

现代工业企业自动化丛书

智能制造概论

曾芬芳 景旭文 等编著
白英彩 主审

清华大学出版社

(京)新登字 158 号

内 容 提 要

本书是《现代工业企业自动化丛书》中的一本。

本书主要介绍了智能制造技术在柔性制造系统和 CIMS 中的重要地位和作用以及智能制造技术的今天与未来。内容包括专家系统的知识表达、获取与推理以及专家系统的开发,神经网络的基本模型和神经网络专家系统,智能制造环境下的产品建模以及智能 CAD 系统的设计,智能化工艺设计 CAPP 系统中零件信息的描述、工艺知识的获取与处理,制造过程中的智能监视、诊断与控制以及柔性装配等制造技术的智能化与自动化等。

本书主要面向工矿企业中的中高级工程技术人员以及管理人员,也可作为工科大学机械及其自动化等专业的教材或参考书。

图书在版编目(CIP)数据

智能制造概论/ 曾芬芳等编著 .—北京:清华大学出版社, 2001
(现代工业企业自动化/ 白英彩主审)
ISBN 7-302-04160-1

 .智... .曾... .计算机辅助制造 .TP391 .73

中国版本图书馆 CIP 数据核字(2000)第 79145 号

出版者: 清华大学出版社 (北京清华大学学研大厦, 邮编 100084)
印刷者: 清华大学印刷厂
发行者: 新华书店总店北京发行所
开 本: 787 × 1092 1/ 16 印张: 16 5 字数: 374 千字
版 次: 2001 年 4 月第 1 版 2001 年 4 月第 1 次印刷
书 号: ISBN 7-302-04160-1/ TP · 2456
印 数: 0001 ~ 4000
定 价: 23 .00 元

《现代工业企业自动化丛书》编委会

名誉主任：张钟俊

顾问：吴钦炜

主任：白英彩

副主任：邵世煌 王行愚 吴启迪 孙廷才

编委：（按姓氏笔划）

于海川 王行愚 白英彩 孙振飞 孙廷才

江志道 刘元元 邵世煌 吴启迪 张兆琪

杨德礼 周德泽 柴天佑 虞孟起 魏庆福

序

当今世界先进工业国家正处于由“工业经济”模式向“信息经济”模式转变的时期,其中技术进步因素起着极为重要的作用,它在经济增长中占 70% ~ 80%。“以高新技术为核心,以信息电子化为手段,提高工业产品附加值”已经成为现代工业企业自动化重要的发展目标。从我国经济发展史来看,其工业经济增长主要是依靠投入大量资金和劳动力来实现的,尚未充分发挥技术进步在工业经济增长中的“二次效益倍增器”的作用。“如何加快发展电子信息技术、调整产业结构,适应世界经济发展需求”是当前我国工业企业自动化界研究的重要课题之一。

工业自动化是一门应用学科,它主要包括单机系统自动化、工业生产过程自动化和工业系统管理自动化等三个方面。企业自动化包括企业生产管理信息电子化、信息处理的自动化以及网络化。现代工业企业自动化涉及到自动化技术、计算机技术、通信技术、先进制造技术和管理学等诸多学科,它需要各学科的专家和工程技术人员通力合作,从而形成“多专业知识与技术集成”的现代工业自动化发展思路。目前工业企业自动化系统主要呈现开放性、集散性、智能性和信息电子化与网络化的特点。在现代工业企业自动化中,计算机控制技术充当了极为重要的角色,它是计算机技术和控制理论有机的结合。自动控制理论的发展是伴随着被控制对象的复杂性、不确定性等因素的研究成果而发展的,它由经典控制理论(频域方法)和现代控制理论(时域方法)发展到第三代控制理论——智能控制理论。计算机控制系统分为数据采集与处理系统、计算机在线操作指导控制系统、直接数字控制系统、监督控制系统、分级控制系统和集散控制系统以及分布式智能控制系统。从当前计算机技术和自动控制技术发展状况来看,高性能工业控制机系统、智能控制系统和基于网络系统的虚拟企业自动化系统将是未来工业企业自动化的重要发展方向。

从系统工程的角度来看,工业自动化技术研究与应用过程分为三个阶段:自动化技术研究阶段、科研成果向实用转化阶段和产品应用阶段。经过我国科技工作者半个世纪卓有成效的研究,在自动化技术研究与应用方面取得可喜的成绩,并给我国的工业自动化事业带来了深刻影响和变革,产生了巨大的社会和经济效益,其中有的技术已经接近或达到世界先进水平,但从应用以及成果向产品的转化的总体发展角度来看,仍存在着一些问题,仍需花大力气进一步探索和研究。例如,我国在工控机及其配套设备的生产方面尚需进一步构成规模经济;建立并发展企业网络及其协议和数据库集成技术,为全面实现我国“金企工程”提供技术和手段;开发系列的工控机软件包、实时操作系统,以提高工控机系统的总体水平;充分运用以工控机为核心的电子信息技术来改造我国各类传统工业的工

装设备及产品;在我国的部分现代企业中大力倡导推行 MIS,MRP- 和 CIMS/ CIPS 以及信息网络系统,以提高企业管理水平和竞争能力等。在 20 世纪 40 年代,计算机刚问世不久,它的应用除在军事、政要部门之外,主要是在各传统工业领域的应用。在 60 年代~70 年代,各国的工业计算机应用极为普遍促进了其工业企业自动化高速发展,而我国的工业企业自动化非但没有大踏步前进,反而停滞不前。到了 90 年代这个问题就显得十分严重了,因此我们必须“补上这一课”。我们编写了《现代工业企业自动化丛书》(目前暂定 42 册,并根据实际需要不断增加新的书目),该《丛书》内容既包括工业生产过程自动化,又包括现代企业管理自动化技术,如基于总线工控机系统、工程数据库、CIMS/ CIPS 以及企业网络技术等。其编写原则为:理论与实践密切结合,为实现工业企业自动化提供典型示范系统。编委会特邀请了国内在该领域有扎实理论基础和富有实践经验的专家分别承担各分册的编审任务,以期在向读者展示国内外相关技术的最新成果和发展动态的同时,提供解决现代企业自动化的思路、方法、技术和设备等。

该《丛书》以工程技术人员为主要读者对象。我们相信该《丛书》的出版必将在推动我国工业企业自动化应用的普及和发展进程中起到积极的作用,为进一步提高我国工业企业自动化水平做出贡献。

清华大学出版社颇具魄力和眼光、高瞻远瞩,及时提出组稿这套《丛书》的任务,他们为编好《丛书》做了认真、细致的准备工作,并为该丛书的出版提供了许多有利的条件,在此深表谢意。同时对于参加各分册编审任务的专家、学者所付出的艰辛劳动表示衷心感谢。编审《丛书》的任务十分繁杂而艰巨,加之时间仓促,书中出现疏漏、欠妥之处也是难免的,希望广大读者不吝赐教,以使我们逐步完善这个《丛书》系列。

中国科学院院士、上海交通大学教授

前 言

制造业是最古老的学科之一。近年来,随着计算机技术的飞速发展,特别是人工智能(AI)技术,如专家系统、人工神经网络等渗透到各个领域,也使制造业发生了巨大的变化。智能制造正是制造技术和人工智能技术结合的产物。

所谓智能制造是在制造生产的各个环节中,以一种高度柔性和高度集成的方式,应用智能制造技术和智能制造系统进行制造的生产模式。智能制造系统是一种由智能机器和人类专家共同组成的人机一体化智能系统,它在制造过程中,通过计算机模拟人类专家的智能活动,进行诸如分析、推理、判断、构思和决策等方式,旨在取代或延伸制造环境中人的部分脑力劳动,并对人类专家的制造智能进行收集、存储、完善、共享、继承和发展;智能制造技术的宗旨在于通过人与智能机器的合作共事,去扩大、延伸和部分地取代人类专家在制造过程中的脑力劳动,以实现制造过程的优化。可见智能制造使制造过程中人的体力劳动通过自动化技术获得了很大的解放,而脑力劳动通过智能决策使自动化程度获得了巨大的提高。智能制造可实现决策自动化的优势使其能很好地与未来制造生产的知识密集型特征相吻合,实现“制造智能”和制造技术的“智能化”。智能制造将是未来制造自动化发展的重要方向。有人预言下一世纪的制造工业将由两个“I”来标识,即 Integration(集成)和 Intelligence(智能)。

本书共分8章。第1章概论,主要介绍了智能制造技术在柔性制造系统和CIMS中的重要地位和作用、智能制造发展的历史背景、国内外智能制造技术的发展状况以及智能制造技术的今天与未来。第2章专家系统,主要介绍了专家系统的知识表达、获取与推理机以及专家系统的开发。第3章神经网络,主要介绍了神经网络的基本模型、学习规则以及基于神经网络的机器学习和神经网络专家系统。第4章智能制造环境下的产品模型及其建模,主要介绍了产品模型的概念,包括几何模型、参数化模型、产品特征模型以及智能制造环境下的产品模型及其建模技术。第5章智能CAD系统及其设计,主要介绍了智能CAD系统的设计模型、设计方法、智能化空间布局设计以及智能CAD系统实例。第6章计算机辅助工艺设计及其智能化,主要介绍了CAPP系统中零件信息的描述、工艺知识的获取与处理以及CAPP专家系统及其研究动向。第7章制造过程的智能监视、诊断与控制,主要介绍了接口技术、智能监视与诊断、制造过程的结构模型、智能控制策略以及制造过程的智能控制等。第8章智能装配,主要介绍了智能装配的装配技术知识表达、装配任务的描述、装配顺序规划、基于特征的装配方法、柔性自动装配系统以及虚拟装配等。

由于智能制造所包含的知识范围十分广阔,同时也由于作者所从事的专业的限制,本书将着重于智能制造技术在机械制造领域中的应用。

本书主要是面向工矿企业的,因此,它适合工矿企业的高、中级工程技术人员和管理

人员使用,也可作为工科大学的教材和研究生的辅导参考资料。

本书由曾芬芳、范牧昌、王建华、俞孟蕪、景旭文等编著,王颖、苏勇、赵丽英等参加了编写工作。全书由白英彩教授担任主审。

智能制造是一个新兴的、发展中的学科,作者衷心希望本书的出版能对理解、推动和开发智能制造系统有所贡献。限于作者的业务水平,疏漏或不妥之处在所难免,敬请读者不吝批评指正。

作 者
2000 年 12 月

目 录

第 1 章	概论	1
1.1	柔性制造系统	2
1.1.1	柔性自动化和柔性制造系统	2
1.1.2	柔性制造系统的基本结构	4
1.1.3	柔性制造系统的实例	5
1.1.4	智能制造技术在柔性制造系统中的应用	7
1.2	计算机集成制造系统	10
1.2.1	计算机集成制造(CIM)和计算机集成制造系统(CIMS)	10
1.2.2	CIMS 的构成	11
1.2.3	智能制造在 CIMS 中的重要地位和作用	12
1.3	智能制造技术的今天与未来	12
1.3.1	智能制造发展的历史背景	12
1.3.2	国内外智能制造技术的发展状况	14
第 2 章	专家系统	16
2.1	专家系统简述	16
2.2	知识的表达与获取	17
2.2.1	知识的分类	17
2.2.2	专家系统的知识表示	17
2.2.3	知识获取	21
2.3	推理机	22
2.3.1	正向推理控制策略	23
2.3.2	逆向推理控制策略	23
2.4	专家系统开发工具	24
2.5	Prolog 语言简介	25
2.5.1	Turbo Prolog 的程序结构	25
2.5.2	Prolog 中的基本编程元素和语法	27
2.5.3	Prolog 的工作方式	30
2.5.4	用 Turbo Prolog 构造一个齿轮设计专家系统	31
2.6	知识工程语言 GEST	33
2.6.1	知识工程语言 GEST 的基本组成和语法	33

2.6.2	事实库及其操作	34
2.6.3	规则库及其操作	36
2.6.4	模式匹配	37
2.6.5	待议事件表及其执行	38
2.6.6	一个简单的应用实例	39
2.7	专家系统的开发	41
2.7.1	开发专家系统的特点	41
2.7.2	开发专家系统的基本步骤	42
2.7.3	专家系统的发展趋势	44
第3章	神经网络	47
3.1	概述	47
3.2	神经网络的基本模型	47
3.2.1	感知机	47
3.2.2	BP 模型	49
3.2.3	Hopfield 模型	51
3.2.4	Boltzmann 模型	53
3.3	神经网络学习规则	55
3.3.1	Hebb 规则	55
3.3.2	Delta 规则	56
3.3.3	ART 规则	57
3.3.4	Kohonen 规则	57
3.3.5	Boltzmann 规则	58
3.4	基于神经网络的机器学习	59
3.4.1	机器学习	59
3.4.2	基于神经网络的机器学习	61
3.5	神经网络专家系统	63
3.5.1	一般描述	63
3.5.2	基于神经网络的专家系统	64
第4章	智能制造环境下的产品模型及其建模	67
4.1	产品模型	67
4.1.1	产品模型的概念	67
4.1.2	产品模型的演变	68
4.2	几何模型	70
4.2.1	几何模型的概念	70
4.2.2	几何模型的建模方法	71
4.2.3	模型的操作	76

4.3	参数化模型.....	77
4.3.1	参数化设计	77
4.3.2	约束	78
4.3.3	参数化建模技术	80
4.4	产品的特征模型及建模.....	88
4.4.1	特征的定义	88
4.4.2	特征的分类	89
4.4.3	特征建模	90
4.5	智能制造环境下的产品模型.....	96
4.5.1	智能制造对产品模型的要求	96
4.5.2	产品的知识模型	98
4.5.3	产品的集成表示模型	99
第5章	智能 CAD 系统及其设计	102
5.1	智能 CAD 系统的设计模型	102
5.1.1	设计模型简介.....	102
5.1.2	基于 CASE 的设计模型	104
5.2	智能 CAD 系统的设计方法	108
5.2.1	面向对象的求解方法.....	108
5.2.2	智能 CAD 的广义类比推理方法	111
5.2.3	设计过程模型的特征设计方法.....	113
5.3	智能化空间布局设计	117
5.3.1	智能化空间布局的分类、建模、求解简介.....	118
5.3.2	三维正交方位图模型及其布局搜索算法.....	123
5.3.3	基于正交方位图与遗传算法的包装布局	127
5.3.4	基于约束满足的多策略布局设计算法及其应用.....	132
5.3.5	基于不同粒度布局中的模糊建模及其应用.....	136
5.4	智能 CAD 系统实例	141
5.4.1	插齿刀智能辅助设计系统.....	141
5.4.2	铣刀设计的智能 CAD 系统	144
5.4.3	机床夹具设计专家系统.....	148
第6章	计算机辅助工艺设计及其智能化.....	151
6.1	简介	151
6.1.1	工艺设计.....	151
6.1.2	工艺设计标准化.....	152
6.1.3	计算机辅助工艺设计(CAPP)发展概述	153
6.1.4	派生式 CAPP	154

6.1.5	创成式 CAPP	156
6.1.6	CAPP 专家系统	158
6.2	CAPP 系统中零件信息的描述	158
6.2.1	零件分类编码法	159
6.2.2	零件的图形要素描述法	160
6.2.3	直接从 CAD 中提取信息	164
6.2.4	特征建模	166
6.3	工艺知识的获取与处理	167
6.3.1	工艺知识的获取	167
6.3.2	工艺知识的表达	170
6.4	应用实例——回转体零件 CAPP 专家系统中的夹具设计	173
6.4.1	夹具设计模块的数据库和规则库	173
6.4.2	实例	178
6.4.3	系统的运行环境	180
6.5	CAPP 专家系统及其研究动向	180
6.5.1	以知识为基础的 CAPP 系统	180
6.5.2	CAPP 开发(工具)环境的研究	181
6.5.3	动态 CAPP	181
6.5.4	面向并行工程的 CAPP	182
第 7 章	制造过程的智能监视、诊断与控制	184
7.1	概述	184
7.2	接口、显示与控制器	185
7.2.1	智能传感器	185
7.2.2	信号变换与接口电路	189
7.2.3	显示与记录设备	198
7.2.4	可编程控制器	201
7.3	智能监视与诊断	205
7.3.1	智能监视与诊断系统的开发	205
7.3.2	系统硬件体系结构设计	207
7.3.3	系统软件设计	210
7.4	制造过程的结构模型	212
7.4.1	输入输出变量的划分	212
7.4.2	结构模型	212
7.5	智能控制策略	213
7.5.1	反馈控制	213
7.5.2	直接数字控制(DDC)	214
7.5.3	最优控制	215

7 5 4	自适应控制.....	216
7 .6	制造过程的智能控制	218
7 6 .1	计算机过程监控.....	218
7 6 2	智能集散型自动控制系统.....	220
第 8 章	智能装配.....	224
8 .1	概述	224
8 .2	人工装配特点	225
8 .3	智能装配	226
8 3 .1	装配技艺知识表达.....	227
8 3 2	装配任务的描述.....	229
8 3 3	装配顺序规划.....	231
8 3 4	基于特征的装配方法.....	236
8 3 5	柔性自动装配系统.....	237
8 3 6	智能装配机器人.....	240
8 3 7	虚拟装配 VA(virtual assembly)	241
参考文献	244

第 1 章 概 论

制造业可能是人类有史以来最古老的产业之一。制造业的发展同工具的使用和发展密切相关,以工具的发明和使用为里程碑,人类经历了石器时代、青铜器时代、铁器时代、蒸汽机时代、电气时代,目前以计算机为工具的信息时代。由于生产工具的不断改进,人类从其所从事的生产环境中不断得到解放,如蒸汽机的发明,大大提高了劳动生产率,使人们有更多的时间从事设计和管理方面的工作。随着制造技术的发展,自动化技术在工业化大生产中得到了迅速的发展,本世纪初,在汽车制造业为代表的大批量生产中,采用凸轮和样板的机床可以高效地生产出大批量相同规格的零件,聪明的设计人员设计出带有自动上下料装置的机床设备并采用传送带把它们连接起来,由于劳动力价格相对便宜,在自动化机床无法使用的地方,采用手工劳动来补充。

随着劳动力价格的不断升高,在大批量生产中节约人力的要求越来越高,可编程控制器和工业机器人得到了广泛的应用,在 20 世纪 80 年代,在汽车制造业中采用点焊机器人和喷漆机器人代替手工劳动,降低了生产成本。

由于社会需求的不断变化,人们要求产品多样化和个性化,市场竞争的加剧,要求产品的更新换代时间不断缩短,这就意味着产品的品种将越来越多,而批量则越来越小。对于加工工业情况也是一样,特别是在飞机制造业、造船工业、机床制造业和模具制造业更是如此,据调查,单件生产和小批量生产比例的增长速度令人吃惊。为了适应这种发展需要,柔性加工技术得到了越来越广泛的应用。近年来,人们对制造过程的自动化给予了极大的热情,资料表明,从 1870 年到 1980 年间,制造过程的效率提高了 20 倍,而生产管理效率只提高了 1.8 ~ 2.2 倍,产品设计的效率只提高了 1.2 倍。以上数据表明体力劳动通过采用自动化技术得到了极大的解放,而脑力劳动的自动化程度则很低。为了进一步解放生产力,特别是提高脑力劳动的自动化程度,柔性制造技术(FMT)和计算机集成制造(CIM)等先进制造技术的研究和开发受到各工业发达国家的普遍重视。

柔性制造技术是近年来国际上一些工业技术比较发达的国家为了进一步提高多品种、中小批量生产自动化,以及提高劳动生产率,降低生产成本,缩短产品的研制或生产周期,增强产品更新换代和市场竞争能力而推出的一种新型制造技术。在柔性制造系统中,机床上装备了工件和刀具的自动更换系统,实现了在线计算机辅助加工过程的调度和规划,出现了完善的加工中心。柔性制造技术至今仍在继续发展之中。

计算机集成制造系统(CIMS)不仅是现有生产模式的计算机化和自动化,而且是在新的生产组织原理和概念指导下形成的一种新型生产实体,也是将在 21 世纪占主导地位的新型生产方式。因此,世界上许多国家和企业都把 CIMS 定为国家或企业的发展战略。

智能制造是人工智能技术和制造技术结合的产物,其目的在实现整个制造过程,即市场预测、生产决策、研究与开发、设计、制造、物流、经营和销售等产品生产的各个环节的自动化和智能化以及它们的集成。它以取代人的部分智能性脑力劳动为目标,以实现制造

过程的自组织能力和制造环境的全面智能化为其最终目标。因此,智能制造技术已在柔性制造系统(FMS)和 CIMS 中得到了广泛的应用。

人工智能理论是智能制造技术的理论基础。目前,国内外对于各种专家系统和人工神经网络的研究为其在制造领域中的应用提供了强有力的技术手段。在本书中的以下两章中对这两方面的基础知识也将加以简单介绍。

1.1 柔性制造系统

1.1.1 柔性自动化和柔性制造系统

在机械制造自动化的早期,由于技术水平的局限,自动化的引入只在大量生产领域里。为了提高劳动生产率,降低生产成本,采用了大批量化、标准化和生产流水线等生产方式,建立起主要由机械式或液压式的自动机床、组合机床或专用机床组成的单一品种的生产自动线。这种自动线有其固定的生产节拍,在这种自动线上,一般无法改变产品的品种,因而也被称为刚性自动化生产线。

到了 60 年代,当国际上多数大批量生产的工厂已实现自动化以后,人们发现,由于市场对商品的需求发生了很大的变化,商品开始向个性化、多样化发展,产品的更新换代周期明显缩短,多品种、小批量生产增加,市场竞争激化。大批量生产只是机械制造业的一小部分,约占 15% 到 25%,而中、小批量生产则占 75% 到 85%,而且其比重还在逐年加大。多品种,中、小批量生产已成为主流生产方式。这就迫使机械制造业的生产企业对市场变化作出快速响应,解决多品种,中、小批量生产的劳动生产率大大落后于大批、大量生产企业的状况。另一方面,由于从 60 年代开始到 70 年代计算机技术和自动化技术得到了飞速的发展,计算机数控(CNC)机床和加工中心在自动化机床领域中逐渐取代了液压式和机械式的自动机床。在各种 CNC 机床和加工中心上,只要改变机床的加工程序就可用来加工各种不同的零件,改变加工对象的灵活性很大,而所需的调整时间却很少,所以 CNC 机床和加工中心非常适合多品种,中、小批量生产领域。由于 CNC 机床本身已具备很好的柔性,故为柔性自动化打下了良好的技术基础。

据统计,在使用普通设备进行多品种小批量生产时,在零件的生产过程中,毛坯和在制品的运输及存放时间上占据了整个生产过程的大部分时间,而花在机床上的时间仅占整个生产过程的一小部分,而在这一小部分时间中,真正用于切削加工的时间一般只占 30% 左右,而 70% 左右的时间用于零件的定位、装夹、换刀、测量,机床的调整以及清除铁屑等。CNC 机床只是减少了机床上部分非切削时间和部分工序间零件搬运存放的时间,因此,为了进一步提高劳动生产率,缩短零件生产周期,充分发挥机床利用率,必须在中、小批量生产领域里进一步提高自动化程度。对上述各个环节尽量实现自动化,如仓库的工件存取自动化、工件输送和搬运自动化、工件装卸自动化、刀具的输送和更换自动化、刀具磨损的自动检测和自动补偿、刀具破损的自动检测与自动换刀、在线检测、自动清除铁屑、机床的运行状态和加工状态的自动监测和自动诊断等。

近年来,由于微电子技术、自动控制技术、计算机网络和数据库技术的迅速发展,以及

微机和可编程控制器在机床、工业机器人、自动仓库、自动导引小车上的广泛应用,通过分布式多级计算机的控制系统就可把它们联结成以数控机床为核心的各种规模的柔性自动化加工方式。

柔性自动化有柔性制造单元、柔性制造系统、柔性生产线、独立制造岛等多种形式。各种不同柔性自动化形式之间的区别不仅在于机器设备方面的配置和复杂程度,还在于信息和数据的处理方法和手段,即信息流和控制系统的特征。按照这种观点来划分,可将不同形式的柔性自动化作如下描述:

柔性制造单元(FMC)通常由1~2台加工中心或其它数控机床组成,并具有不同形式的刀具交换和工件的装卸、输送及存储功能。除了机床的数控装置外,通常还有一台单元计算机来进行数控程序的管理和外围设备的协调。

柔性制造系统由两台以上的加工中心或柔性制造单元以及清洗、检测设备组成,具有较完善的刀具和工件输送和存储系统,除主控计算机(又称单元控制器)外,下面还设有工作站级和设备级,形成三级控制组成的局部网络。

在柔性制造系统中由于实现了较高水平的自动化和智能化,可大大提高生产效率,有许多系统还实现了日班把工作准备好,夜班无人化加工的生产方式,进一步提高了劳动生产率。

柔性制造系统的适用范围主要取决于其制造批量,图1.1所示为各种不同加工方式的加工范围。图中纵坐标为各种加工方式适用的批量加工零件数,横坐标表示适用的加工零件种类数。从图中可以看出,柔性制造系统的适用批量加工零件数是50~2000件,而适用的加工零件种类数是4~100种。

图 1.1 制造方式与零件种类和批量的关系

以FMS为代表的柔性自动化生产经过70年代的研制和实验阶段,到80年代已进入实用和推广阶段。近年来,其发展方向主要有两个:一个方向是向大型化、高度自动化、智能化的柔性制造工厂(FMF)方向发展。建造FMF的企业一般在技术上和经济上均需有雄厚的实力,这是因为FMF技术密集、投资巨大、建造周期长(5~7年),投资回收

的时间也较长,其技术经济效益很高。另一个方向是向廉价化、小型化的柔性制造单元方向发展。由于 FMS 耗资大、技术复杂,因此许多中小企业采用分阶段实现 FMS 的方法,采用的方法是首先建立 FMC 甚至更小规模的单机柔性制造单元(也有称为柔性制造模块 FMM),等到取得经验后再逐步扩大,这种方法称为“逐步扩展法”或“增量法”。为适应这一潮流,许多机床厂家把 CNC 机床设计成能适用于可扩展为 FMC 和 FMS 的具有高度自动化的加工设备,这种设备带有交换工作台、自动换刀机械手和一定数量的刀库、在线工件测量和自动排屑装置等,可以进行无人化加工。为了方便使用,机床设计实现模块化,既可单独使用也可纳入 FMC 和 FMS。这些机床一般带有串行接口和 DNC 接口,通过这些接口可以很方便地和上级计算机相连。

柔性较小的,具有自动生产线特征的柔性制造系统称为柔性生产线。

独立制造岛(AMI)是我国同济大学张曙教授首先提出的一种柔性制造技术。独立制造岛是以几台 CNC 机床和数控加工中心为主体,同时配备有普通机床和手工操作工位(如毛坯准备和检验站),不同形式的物料输送装置以及相应的工件库、刀具库和夹具库,在计算机软件的支持下组成能完成一定范围零件加工的制造系统。独立制造岛的机床布置,不同于传统的机群方式,而是按成组技术的原理,转变为加工过程相对封闭的单元,连同工件库、刀具库和夹具库,构成了独立制造岛的硬件基础。独立制造岛的软件采用模块化结构,分为计算机辅助工艺设计子系统、数控程序管理子系统、计算机作业计划和调度子系统、物料管理子系统和质量控制子系统等五个子系统。各子系统之间及系统内部通过数据接口进行信息交换,软件系统设有公共数据库,数据库内存有零件、资源、工艺及工况等方面信息供各子系统共享,并可通过局域网与上级计算机管理系统相连,这样就构成了独立制造岛的信息基础。独立制造岛结合我国一部分中小企业资金有限,无法大量购买先进制造设备的具体情况,在引进部分关键数控设备的同时,允许采用部分普通设备,不过分强调无人化生产和完全自动化,而是把重点放在先进管理技术、软件开发方面,强调人的因素。事实证明,独立制造岛是一种适应我国国情的准柔性制造系统。

1 1 2 柔性制造系统的基本结构

柔性制造系统一般由 CNC 加工设备、物流系统和计算机控制系统三部分组成。

1 . CNC 加工设备

柔性制造系统根据其加工对象的不同,可分为以加工箱体类零件、回转体零件、混合型零件、钣金零件等为主的加工系统。对于以加工箱体类零件为主的柔性制造系统,一般配备有多台立式和卧式加工中心。对于以加工回转体零件为主的柔性制造系统多数配有 CNC 车削中心或 CNC 车床,由于许多回转体零件上还有平面或键槽加工,所以这类系统中往往还配备有立式或卧式加工中心。对于一些专门零件加工,如齿轮加工,还需要配备 CNC 齿轮加工机床等。

在加工较复杂零件的柔性制造系统中,加工中心本身的刀库容量往往不能满足加工的需要,因此,在许多柔性制造系统中都设有自动刀库以补充机载刀库容量的不足。

2 . 物流系统

柔性制造系统的物流系统与传统的自动线或流水线的零件传送系统有很大差别,柔性制造系统的零件传送没有固定节拍,也没有固定顺序,有时甚至是几种零件混杂在一起传送,也就是说,整个物流系统的工作状态是可以随机控制的。

柔性制造系统的物流系统可以有多种型式,对于箱体类零件,工件经常装在托盘上进行输送和搬运。该系统包括工件在机床之间、加工单元之间、自动仓库和托盘存放站之间,托盘存放站和机床之间的输送和搬运等。自动搬运设备有链式传送带、滚筒式传送带、有轨小车、无轨自动导引小车、悬挂式机械手等;托盘存放站与机床之间的装卸设备有托盘交换台 (APC)和机器人,装卸工件的机器人又可分为内装式机器人、附装式机器人和单置万能式机器人等,在柔性制造系统中,自动导引小车有时也用来输送刀具。

3 . 计算机控制系统

根据柔性制造系统的规模大小,控制系统的复杂程度也有所不同,通常多为三级分布式控制系统,低级为控制级,高级为决策级。在各级的决策与控制中,生产的计划与调度、加工过程途径的确定是主要问题。第一级为过程控制及逻辑控制级,其主要功能是对加工设备和工件装卸机器人或托盘交换台的控制,包括对各种加工作业的控制和监测等,其计划时限 (planning horizon)为数毫秒至数分钟。第二级为工作站控制级,其主要功能是对柔性制造系统中各种自动化环节或 FMS 分系统进行控制与管理。其对象包括物流运送、自动料仓存取、刀具管理、清洗、在线检测及自动加工单元等,其计划时限为数分钟至数小时。第三级为单元控制级,其主要功能是对各工作站进行管理和控制,因此有时也称为单元控制器 (cell controller),这一级控制主要负责生产管理,编制日程进度计划,把生产所需的信息,如加工零件的种类和数量、每批的生产期限、刀夹具种类和数量等送到第二级管理计算机。单元控制级的计划时限为几小时至几十小时。

1 1 3 柔性制造系统的实例

由于国际市场上柴油机的需求的变化,使日本的 Diesel Kiki 柴油机制造公司决定改变其生产方式,从传统的大量生产方式转到柔性制造方式。

Diesel Kiki 柴油机制造公司原来依靠传统的专用机床和自动生产线制造喷油泵体,这种生产方式适用于大批量生产。但随着市场需求的不断变化,产品品种的需求量大大增加,因而需要从成品库控制的生产方式改变为面向用户的及时供应产品的生产方式,为此,有必要采用柔性制造系统进行生产。

Diesel Kiki 公司的柔性制造系统由 3 台 Yasda 卧式加工中心和一台 Yasda 立式加工中心、一个单层托盘库(容量为 70 个托盘)、一台带自动托盘交换装置的有轨小车及一个装卸站组成,系统由一台微机控制。用一个集中的切屑与冷却液系统和传送带排除切屑。

加工 8 种泵体的全部零件的 NC 程序存储在各台机床的控制器内,使得每台机床可以在系统内单独地工作。Yasda YBM-70 卧式加工中心具有 240 KB 存储器和一个 120

把刀的刀库。YBC-80 V 立式加工中心具有 128 KB 存储器和一个 40 把刀的刀库。

图 1.2 为 Diesel Kiki 公司的 FMS 布置图。图中各数字代号分别说明如下:1 .托盘库(70 个工位),2 .装卸站,3 .系统控制器,4 .有轨小车控制器,5 .有轨小车,6 .YPC-80V-50P 立式加工中心,7 .YBM-70N-50AP 卧式加工中心三台,8 扩展的 YBM-70N-50AP 一台,9 .预加工机床,10 .专用钻床,11 .插床,12 .切屑与冷却液站,13 .对刀仪,14 .空气压缩机。

图 1.2 Diesel Kiki 公司的 FMS 布置图

柔性加工系统的工作情况为:

在每个零件进入柔性加工系统以前,先钻三个孔,攻丝并在一边刨平面。这样可以使得工件在系统内的加工过程中只需一次装夹。在一台卧式加工中心上可加工零件的四个面,而顶面则在立式加工中心上加工。

在两个日班的每一班有两个操作人员看管系统的运行,而夜班则无人。操作人员负责预加工、在托盘上装卸零件和工具的调定。他们还负责后两个加工工序:去毛刺和检验。

在装卸站,待加工的零件借助于三个螺钉,穿过三个预先钻好的孔,安装在托盘夹具上。托盘夹具按成组夹具的思想设计,可适用在系统内所有被加工的零件上。人工安装时间约为每个零件两分钟。借助有轨小车将零件自动地运送到托盘库。

在计算机的控制下,由有轨小车把工件托盘从托盘库输送到一台加工中心的托盘交换装置,它已经发出了信号准备接受新的工作。由气动装置把托盘自动夹紧在加工中心的工作台上。在不同机床上的加工工序之间,工件托盘往返回到托盘库。正常情况下,工件按先进先出的原则从托盘库调出。

托盘输送系统、托盘库、装卸站和机床等由微机统一管理,各环节之间采用 RS-232 接口进行连接,以形成一个自动化生产系统。

进入系统的每个工件都带有工作序号,这些工作序号是由操作人员在装卸站通过键盘输入的。从此以后,由储存在微机内的程序对物流进行控制。这种控制把机床编号与程序编号和已知工件编号关联起来,并提供一个可选择的机床编号。控制计算机将根据来自机床、输送系统和装卸站的实时数据不断地修正和更新物流的生产日程进度计划。

Diesel Kiki 公司的喷油泵壳体柔性制造系统的总投资约为 500 万英镑,投产后取得以下效益:

缩短生产周期;
减少库存和在制品;
简化生产控制;
能够 24 小时连续生产;
大幅度减少了人工干预;
在产品品种与数量变化的情况下,能保证稳定的生产率;
通过减少零件的装夹次数,提高了产品质量;
使生产具有柔性,能适应公司经营产品的变化。

投产一年后,该系统已能生产 8 种不同类型的铝制泵体和铸铁泵体。在卧式加工中心上的加工循环时间为 50 分钟,在立式加工中心上的加工循环时间为 12 分钟。系统采用三班工作制,月产 2000 个喷油泵体。系统采用混流方式生产,由装配的需要决定加工的零件。系统的机床利用率在 20 ~ 22 小时/日之间,该公司计划在两年内收回投资。

1 1 4 智能制造技术在柔性制造系统中的应用

1. 柔性制造系统的规划决策专家系统

由于柔性制造系统投资庞大,建设的周期长,投资有一定的风险,因此必须进行认真的规划。在规划过程中,最重要的是要使设计中的柔性制造系统满足预计的各项性能指标,并防止系统的规模设计得过高和过低,避免浪费资金和导致系统规划的失败。一般系统的规划设计主要可分为以下三个步骤:

(1) 零件族和机床的选择 首先根据成组技术选择一族适合于柔性制造系统加工的零件,并根据零件的加工工艺选择最经济有效的机床、刀具和夹具。

(2) 总体方案的设计 估算零件的批量和每批数量,估算各类机床的最少台数,制定各种不同的系统方案,其中包括物流系统和其它辅助设备的设计。

(3) 分析、评审总体方案 模拟各种规划设计方案,分析系统的性能指标,改进系统的设计方案,审定难以确定的指标如柔性、精度等。

在上述的每一步骤间都必须加以反馈分析和经济分析。此外,由于柔性制造系统是一个随机的、离散的加工系统,因此系统的规划设计与将来系统的运行,即系统的作业计划调度等又是不可分割的,然而在系统的负荷优化和最佳生产控制方式等方面都是目前难以解决的问题。至今还没有一套完整的数学模型,如采用人的经验就需要花费大量的时间和费用来熟悉实际系统的运行规律,所以解决这一问题的一个理想途径是将计算机的运算能力和人的实际经验相结合,即采用人工智能的方法来建立相应的专家系统。

采用专家系统来实现柔性制造系统的规划决策仿真系统,系统由仿真模型和专家系统两大部分组成,两部分之间通过接口相联。仿真系统输出的报告通过接口转化成专家系统能利用的知识信息,而专家系统给出的建议则通过接口转化为仿真系统能接受的参数集。

采用上述方法由于需要将仿真算法和专家系统的知识信息组合起来,因而使仿真系统难以修改和扩展。为了解决上述问题,有人提出一种面向对象的仿真系统设计方法,旨

在实现更多基于知识的推理来取代参数模型,这种思想为 FMS 的智能化仿真方法开辟了一条新路。

2 . 智能化设备

在一些比较先进的柔性制造系统中,目前已采用或正在研制智能化设备。据有关文献介绍,这些智能设备包括智能化自动仓库、智能机器人、智能加工中心、带有智能控制的自动导引小车等柔性制造系统中的关键设备。

智能化设备一般由感知部分、智能控制部分和执行部分组成。

感知部分主要由各种感受器组成,其中应用最多的是光感受器和广义力感受器。在柔性制造系统的智能化设备中,光感受器主要用于工件尺寸检测、几何位置和姿态的检测、运动监视、形状识别(工件识别、刀具识别)、表面粗糙度检测、加工过程检测、加工设备状态检测、环境监视等。而力感受器则多用于机器人和智能测量仪的测头等方面。

智能控制部分主要由专家系统和控制部分组成,由各种传感器得到的信号通过信息转换变成专家系统可以接受的符号信息,然后利用专家系统中的推理机进行推理,决定要采取的行动。

执行部分的功能主要是搬运、装配、装卸工件、测量和完成其它各种工艺作业。

3 . 柔性制造系统的智能化生产调度

柔性制造系统的生产调度是指系统的单元控制器根据生产计划,在一定期间内给一系列生产设备分配任务,对作业者或不同的设备规定作业的开始、结束的时间或作业的顺序。通过对流动在各台设备(加工中心、物料小车、清洗机、装卸站等)之间的每个工件的任务进行调配,完成其生产调度任务。

柔性制造系统生产调度是 FMS 实际运行中一个不可缺少的重要环节,合理制定 FMS 动态调度的优化策略,对提高系统的生产效益具有十分重要的意义,但如果采用将调度决策规则固化在程序中的方法,则存在决策规则可变性差、设计的软件通用性差、FMS 调度的柔性不足等缺点。

在 FMS 生产调度系统中如果采用专家系统,在系统中用推理机的推理功能代替线性、非线性规划中的复杂计算,可以使调度系统满足复杂系统的实时性要求。专家系统由工作数据库、知识库、推理机和人机接口组成,其中数据库中存放有关系统的静态、动态描述;而知识库和推理机则随知识的表达方式和推理方式的不同而不同。

用户对 FMS 调度的目标要求不尽相同,但总的来说不外乎满足交货期、保证尽可能高的生产率和减少在制品等几个方面,而这些指标的实现主要是通过较好的作业分组和排序、解决瓶颈设备(如运输小车和清洗设备)的冲突、合理选择机床等主要途径。这里要考虑的主要因素包括各机床的功能(如可进行作业的类型)和状态(如作业列队)、各种作业本身的特性(如所需刀具类型、加工时间)、刀具分布(如现有加工中心中刀具的种类和数量)、瓶颈设备的状态(如小车运动列队长度和位置、各任务的申请先后和位置及目标位置)等。它们是确立知识表示模式和建立推理机的直接依据。

4. 柔性制造系统的智能生产过程监控系统

在柔性加工系统中,为了节约人力,往往采用夜班无人化运转方式,因此,FMS 的生产监控系统是保证系统高效可靠工作的关键。在 FMS 的实时生产控制系统中,对于刀具的监控和对于工件的监控是两个最主要的部分。

对刀具的监控主要指 FMS 中各机床刀库中各刀具的地址、刀具剩余寿命、姐妹刀的情况,在 FMS 中央刀库中各刀具的地址和使用寿命情况。在 FMS 中刀具的动态流动情况等的监控,包括刀具上下主轴、机床刀库换刀、中央刀库与机床刀库中的刀具交换等刀具状态的监控及更换损坏刀具的监视。刀具使用寿命的监视及更换寿命已到刀具的监视等。

对工件的监视主要指在托盘上进入 FMS 中所有工件的品种、数量、工序和地址的监视;在加工中心对工件加工质量的监视及对不合格工件的监视。

在以上监控内容中,刀具状态的监控是加工过程监控的一个十分重要的内容,刀具状态对加工质量的影响存在于整个加工过程中,而且刀具的突发事件,如刀具的意外破损、钻头折断等的概率较高,因此,必须对刀具的状态进行实时监控。刀具的状态信息是通过传感器获得的,刀具状态的监控中常用的传感器有固体介质 AE(声发射)传感器、功率信号监控器、切削振动信号传感器等。由于切削加工的多样性和切削参数的多变性,目前刀具状态监控的信息采集有向多传感器化、特征提取的多模型化发展的趋势。

刀具磨损的一般过程包括信号检测与处理、特征提取、决策与控制几部分。可实现智能监控的主要策略有:专家系统、模糊判别、人工神经网络等。在刀具的实时监控方面,人工神经网络作为一种多输入、多输出的非线性系统,其联想、记忆、抽象、自适应的特征,对解决不确定性和非线性较大的刀具监测问题具有很大的潜力。

由于刀具的磨损是逐步发生的,故在一定切削条件下测得某一磨损状态时的各种相应信号(如声发射、切削振动、主轴功率等)是不难实现的。将传感器信号作为网络输入,而与其对应的磨损状态作为期望输出,就可以获得足够多的实际范例。以此作为训练样本,通过网络各层结构,利用网络实际输出和训练样本所提供的期望输出值调整网络内部表达(主要指权值)来实现自适应地学习,而无需对训练样本的概率分布作过多的要求,便可最终达到正确的分类。

据文献资料介绍,目前已有很多研究机关和院校成功地利用人工神经网络对刀具状态进行识别,实验结果证明,人工神经网络对刀具磨损状况有很好的识别能力。

5. 柔性制造系统的故障诊断

柔性制造系统使多品种、中小批量生产实现了高度自动化,但在系统的高度自动化的同时也造成系统结构的复杂性,这给故障的诊断和排除带来很大困难,它已成为柔性制造技术发展的关键问题之一。

一般来说,柔性制造系统的故障有以下几个特点:

FMS 运行时,其系统行为、状态及部件之间存在着复杂的关联性,某一设备的故障可能会影响到其它设备的运行;

多种故障可能并发,需要判别故障的影响程度,优先处理最紧急的故障;

系统故障要求在线自动采集信息,实行实时诊断和处理。

从上述特点可以看出,柔性制造系统的故障诊断要比一般单一设备的故障诊断复杂得多,其故障诊断与处理涉及到多种学科知识,是一种复杂的人类智能活动。传统的诊断方法已很难满足系统的实时性要求,采用以知识为基础的专家系统,可以比较好地解决这一问题。

在柔性制造系统的故障诊断系统中,知识按内容可分为以下三类:

第一类是故障描述知识,其主要内容是信号的格式与含义说明。系统的故障信息可由故障信号检测装置直接提供,并对这些故障信息规定了具体的表达格式,该表达式包含的故障信息主要有:

故障检测装置编号,它表明故障信号的来源地;

故障级别号,用来表明故障类别,即故障要求处理的优先等级;

故障编号,故障编号对应一种实际故障。

第二类为故障的监控与诊断知识,其主要内容是从领域专家那里得到的关于系统的监控与诊断的知识。

第三类为故障处理知识,该部分是领域专家的关于故障处理以及 FMS 的运行和控制等有关方面的知识。

在故障诊断系统中,其数据库中的内容可以是故障发生时检测到的数据存储,也可以是人为检测的一些特征数据。知识库中存放的一般知识可以是故障的各种现象,故障产生的原因,故障产生时所对应的子系统、设备、部件。规则集是一组规则,反应系统的因果关系,如故障现象与部件的对应关系、故障部件与故障模式的对应关系等。人机接口主要用来为数据库和知识库提供故障前和故障发生时观测到的新现象、故障的新原因、故障发生的新规则等,它提供了知识获取、修改、扩充和完善的维护手段。推理机控制并执行故障原因的求解过程。

1 2 计算机集成制造系统

1 2 1 计算机集成制造(CIM)和计算机集成制造系统(CIMS)

在讲 CIMS 以前我们首先来看一下什么是 CIM。

CIM 直译为计算机集成制造 (computer integrated manufacturing)。它是美国约瑟夫·哈林顿(Josph .Harrington) 博士针对企业所面临的激烈市场竞争形势而提出来的组织企业生产的一种哲理。此后,CIM 在世界各工业国的推动下,20 多年来,特别是近十多年来,CIM 概念不断得以发展,但至今尚无权威性定义,在本书中,我们选用几个比较权威的定义:

日本能率协会在 1991 年完成的研究报告中提出,CIM 的定义为:“为实现企业适应今后企业环境的经营战略,有必要从销售市场开始对开发、生产、物流、服务进行整体优化组合。CIM 是以信息作为媒介,用计算机把企业活动中多种业务领域及其职能集成起

来,追求整体效率的新型生产系统。”

欧共体 CIM-OSA 汇报课题委员会(由 21 个欧洲自动化公司和研究机构,包括总部在美国的 IBM,DEC,HP 公司在欧洲的分公司)最近提出了一个比较权威的定义:“ CIM 是信息技术和生产技术的综合应用,旨在提高生产型企业的生产率和响应能力,由此,企业所有的功能、信息、组织管理方面都是一个集成起来的整体的各个部分。”

我国 863 计划 CIMS 主题专家组组织编写的《计算机集成制造系统(CIMS)约定、标准与实施指南》一书中给 CIM 的定义是:“ CIM 是一种组织、管理与运行企业生产的新哲理,它借助计算机硬软件,综合运用现代管理技术、制造技术、信息技术、自动化技术、系统工程技术,将企业生产全部过程中有关人、技术、经营管理三要素及其信息流与物流有机地集成并优化运行,以实现产品高质量、低消耗、上市快,从而使企业赢得市场竞争。”

如前所述,如果说 CIM 是组织现代化生产的一种哲理,一种指导思想,那么 CIMS 便是这种哲理的实现。

1 2 2 CIMS 的构成

CIMS 通常是由管理信息系统、产品设计与工程设计自动化系统、制造自动化系统、质量保证系统、计算机网络和数据库系统六个部分有机地集成起来的,所以一般说 CIMS 是由四个功能分系统和两个支撑分系统组成。但这并不意味着实践中任何一个企业、工厂实施 CIMS 都必须实现这六个分系统,而应根据具体需求、条件,在 CIM 思想指导下局部实施或分步实施。

下面对六个分系统做简要介绍:

(1) 管理信息系统的功能和构成 管理信息系统是以制造资源计划 MRP (manufacturing resource planning)为核心,包括预测、经营决策、各级生产计划、生产技术准备、销售、供应、财务、成本、设备、工具、人力资源等管理信息功能,通过信息的集成,达到缩短产品生产周期、降低流动资产占用、提高企业应变能力的目的。

(2) 产品设计与工程设计自动化系统的功能和构成 产品设计与工程设计自动化系统是用计算机来辅助产品设计、制造及产品性能测试等阶段的工作,即常说的 CAD/ CAPP/ CAM系统,目的是使产品开发活动更高效、更优质、更自动地进行。

(3) 制造自动化系统的功能和构成 制造自动化系统是 CIMS 中信息流和物料流的结合点,它可以由数控机床、加工中心、清洗机、测量机、运输小车、立体仓库、多级分布式控制计算机及相应的控制软件构成。

(4) 质量保证系统的功能和构成 质量保证系统包括质量决策、质量检测与数据采集、质量评价、控制与跟踪等功能。系统保证从产品设计、制造、检验到售后服务的整个过程,以实现产品的高质量、低成本、提高企业的竞争力的目的。

(5) CIMS 的数据库分系统 从 CIMS 看,数据管理系统和数据库管理系统是一种支撑系统,它是信息集成的关键之一。CIMS 环境下的经营管理信息、产品设计与工程设计自动化、制造自动化、质量保证等四个功能系统的信息数据都要在一个结构合理的数据库系统里进行存储和调用,以满足各系统信息交换和共享。

(6) CIMS 的计算机网络系统 计算机网络技术是信息集成的关键技术之一。计

算机通信网络要提供系统互连和信息互通的能力,通过通信网络将物理上分布的 CIMS 四个子系统的信息联系起来,达到信息共享的目的。

依照企业覆盖地理范围的大小,有两种网络可以选用,即局域网和广域网,如果工厂厂区的地理范围不十分大,一般以互连的局域网为主。

当前,局域网采用国际标准化组织/开放系统互连(ISO/OSI)协议作为实际的工业标准,低层选用电气电子工程协会(IEEE)所制定的 802.X 协议。这种网络成本低,性能高,站点接入和拆除都比较方便。它适合于厂区以信息管理为特征的 CIMS 环境。在比较大型的工厂实现 CIMS 时,可能需要把与多个自动化孤岛相应的多个不同的计算机网络互连起来,通常采用 TCP/IP 事实上的标准实现这些异常网络互连。

1 2 3 智能制造在 CIMS 中的重要地位和作用

从目前来看,有关智能制造(IM)的研究和开发的内容都和 CIMS 密切相关,CIMS 的四个主要分系统,即管理信息系统、产品设计与工程设计自动化系统、制造自动化系统、质量保证系统中的许多问题都可利用智能制造技术得到解决,如管理信息系统中的市场预测系统、经营决策系统和各级生产计划系统都需要大量的专家知识。在产品设计与工程设计自动化系统中,目前正在大力开展智能 CAD 和 CAPP 专家系统的研究,从而进一步提高设计和工艺设计自动化水平。智能制造技术在制造自动化系统中已得到广泛的应用,如前面所提到的智能化设备、柔性制造系统的故障诊断专家系统、智能生产过程监控系统、生产调度专家系统等。在质量保证系统中,智能制造技术也得到了广泛的应用,如一些智能检测和试验装置等。

综上所述,我们认为,在 CIMS 工程的开发过程中,人工智能技术已经发挥了很大的作用,随着 CIMS 技术在我国为进一步推广应用,必将推动智能制造技术的进一步发展。专家们预言,21 世纪将进入智能自动化时代。

1 3 智能制造技术的今天与未来

1 3 1 智能制造发展的历史背景

当历史步入 20 世纪 60 年代后,由于市场经济的冲击和信息革命的推动,作为世界财富主要来源的制造业正在经历一场重大的变革。

首先从市场竞争来看,企业面临的是一个变化越来越快的市场和越来越激烈的竞争环境,社会对产品的需求正从大批量产品转向多品种、小批量甚至单件产品上。企业要在这样的市场环境中立于不败之地,必须对自身不断进行改造以适应变化了的的市场。柔性制造系统和计算机集成制造系统正是为了适应以上市场变化的新型制造系统。

在柔性制造系统中,实现了工件和刀具的自动更换,实现了计算机在线的机床加工过程调度和规划,出现了完善的加工中心。采用柔性制造技术,大大提高了企业在生产中、小批量产品时的自动化水平,目前柔性制造技术仍处于不断发展和完善的阶段。

和柔性制造系统相比,计算机集成制造系统则是一种更大规模的制造系统。根据计

计算机集成制造的哲理,制造系统同时也是一种信息系统,在制造业过程信息处理就是对市场信息、开发信息、制造信息、服务信息和管理信息的获取、加工和处理的过程。最终所得的产品实质是物质、能量和信息三者的统一体。根据这一哲理,计算机集成系统强调的是企业内部信息的集成和信息的共享,从而实现从销售市场开始对开发、生产、物流、信息流服务进行整体优化组合,进而提高生产型企业的生产率和响应能力。

人工智能(AI)技术自 1956 年问世以来,在 AI 研究者们的努力下,40 年来无论在理论和实践方面都取得了重大的进展,特别是 1965 年 Stanford 大学计算机系的 Feigenbaum 提出为了使人工智能走向实用化,必须把模仿人类思维规律的解题策略与大量的专门知识相结合。基于这种思想,他与遗传学家 J .Iederberg,物理化学家 C .Djerassi 等人合作研制出了根据化合物分子式及其质谱数据帮助化学家推断的计算机程序系统 DENDRAL。此系统获得极大成功,解决问题的能力已达到专家水平,某些方面甚至超过同领域的专家。DENDRAL 系统的出现,标志 AI 的一个新的研究领域——专家系统的诞生。

随着专家系统的成熟和发展,其应用领域迅速扩大,70 年代中期以前的专家系统多属于数据信号解释型和故障诊断型,70 年代以后专家系统的应用开始扩展到其它领域,如设计、规划、预测、监视、控制等各个领域。

神经网络是人工智能的另一个重要发展领域,特别是 1987 年 IEEE 召开了第一次国际神经网络会议后,神经网络的理论与应用的研究进入了一个蓬勃发展的新阶段。迄今,神经网络的研究已获得诸多方面的新进展和新成果:提出了大量的网络模型,发展了许多学习算法,对神经网络的系统理论和实现方法进行了成功的探讨和实验。在此基础上,人工神经网络还在模式分类、机器视觉、机器听觉、智能计算、机器人控制、故障诊断、信号处理、组合优化问题求解、联想记忆、编码理论和经营决策等许多领域获得了卓有成效的应用。

计算机技术自从问世以后,迅速在制造业中得到广泛的应用,在软件方面,有计算机辅助设计(CAD)、计算机辅助工艺设计(CAPP)、计算机辅助制造(CAM)、管理信息系统(MIS)、制造资源计划(MRP)、数据库等大量计算机辅助软件产品。在硬件方面有计算机数控机床、工业机器人、三坐标测量仪和大量的由计算机或可编程控制器进行控制的高度自动化设备。上述软、硬件和计算机网络技术的发展,为柔性制造系统、计算机集成制造系统等先进制造技术提供了基本的技术支撑。

柔性制造系统和计算机制造系统的广泛应用不但大大提高了制造业的自动化水平,而且为各种人工智能技术在制造领域的应用提出了更为迫切的需求和更为有利的实施环境。从目前发表的大量文献资料来看,目前人工智能在制造业的研究和应用,特别是在 CIMS 中的应用主要集中在经营决策、生产规划、制造加工和质量保证等几个方面,首先在各 CIMS 的各子系统内分别实现了智能化。

将人工智能技术引入制造领域,对于制造业来说无疑是一场革命性的变革。它的出现将使人们从一个完全崭新的角度去从事制造领域的科学研究,其研究结果对于 FMS, NC, CNC 乃至一般工业过程自动化和精密生产有很大的实际应用价值。它将不仅促进 CIMS 达到更高程度的集成,而且将导致一种新型的制造系统——智能制造系统(IMS)

的诞生。科学家们断言:制造业的发展道路必将是: NC—FMS—CIMS—IMS 。显而易见,智能制造是面向 21 世纪的制造技术和重大研究课题,是现代制造技术、计算机科学与人工智能发展的必然结果,也是三者结合的产物。

1 3 2 国内外智能制造技术的发展状况

近年来,西方各国对智能制造技术的研究和开发都十分重视。美国是智能制造思想的发源地之一,美国政府高度重视智能制造,将其视为 21 世纪占领世界制造技术领先地位的基石。1991 ~ 1992 年度和 1992 ~ 1993 年度,美国国家科学基金(NSF)着重资助了有关智能制造的诸项研究,这些项目覆盖了智能制造领域的绝大部分。美国还建立了许多重要试验基地,美国国家标准和技术研究所的自动化制造与试验基地(AMRF)就把“为下一代以知识库为基础的自动化制造系统提供研究与实验设施”作为其中目的之一。卡内基梅隆大学(CMU)的制造系统构造实验室一直从事制造智能化的研究,包括制造组织描述语言、制造知识表示、制造通讯协议、谈判策略和分布式知识库,先后开发了车间调度系统(ISIS)、项目管理系统(CALLISTO)等项目。在美国空军科学制造计划的支持下,于 1989 年由 D .A .Boume 组织完成了首台智能加工工作站(IMW) 的样机。该样机能直接根据零件的定义数据完成零件的全自动加工,具有产品三维实体建模、创成式工艺规划设计、NC 程序自动生成、加工过程智能监控等一系列智能功能,它的完成被认为是智能制造机器发展史上的一个重要里程碑。与此同时,美国工业界也以极高的热情投入智能制造的研究开发,1993 年 4 月在美国底特律由美国工程师协会召开的 IPC '93(22 届可编程控制国际会议),有 200 多家厂商参展,以极大的篇幅介绍了智能制造,提出了“智能制造,新技术、新市场、新动力”的口号,展出了大量先进的、具有一定智能的硬件设备。这次大会讨论的议题有开放式 PLC 体系及标准、模糊逻辑、人工神经网络、自动化加工的用户接口、通往智能制造之路、精良生产等。

日本由于其制造业面临劳动力资源短缺、制造产业向海外转移的空洞化,制造技术高度的内部化导致标准规范不统一以及与美欧贸易摩擦加剧等问题,迫切感到发展智能制造的重要性。日本东京大学 Furkawa 教授等人正式提出智能制造系统(IMS)国际合作计划,并于 1990 年被日本通产省立案为国际共同研究开发项目。欧洲共同体委员会、日本通产省、美国商务部于 1990 年 5 月经协商成立 IMS 国际委员会,以 10 年为期限,投资 1500 亿日元,实验研究智能制造系统。日本提出的智能制造系统国际合作计划以高新计算机为后盾,其特点是:

- 由政府出面支持,投资巨大,民间企业热情高,有 50 多家民间企业参加;
- 强调部分代替人的智能活动,实现部分人的技能;
- 使用先进的智能计算机技术,实现设计、制造一体化,以虚拟现实技术实现虚拟制造;
- 强调全球性网络制造的生产制造;
- 强调智能化、自律化智能加工系统和智能化 CNC、智能机器人的研究;
- 重视分布式人工智能技术的应用,强调自律协作代替集中递阶控制。

目前,国际上十分重视智能制造技术的研究,专家们预言,21 世纪将进入智能自动

化时代。

我国在智能制造的研究方面起步较晚,但在这个领域的研究发展还是比较快的,从发表的文献资料来看,目前绝大多数的研究集中在人工智能在制造业的各个领域的应用方面,如智能 CAD、CAPP 专家系统、机电设备的智能控制、加工过程的智能检测、智能化生产调度系统、生产决策系统等,许多研究成果已经在实际生产中发挥了很大的作用。

1989 年在华中理工大学召开的“机械制造走向 2000 年——回顾、展望与对策”大会,云集了一大批机械学科的著名专家学者,有不少专家学者就 AI 在制造领域中的应用进行了探讨,并首次把智能制造系统 (IMS) 提到议事日程上来。1990 年,华中理工大学首次组建了 IM 学科组,积极跟踪国际 IMS 的最新研究动态和从事 IMS 关键技术的预研工作。1991 年,杨叔子教授等人首次就智能制造技术和智能制造系统进行了客观的评价,指出它是面向 21 世纪的制造技术,并提出了我国近期研究的多项关键技术,这对我国机械学科的改造与发展及“九五”科技攻关立项无疑有十分重大而又深远的意义。我国国家自然科学基金会从 1993 年起每年都有一定的资金用于资助智能制造方面的研究项目。我国的“九五”规划中已把先进制造技术(包括智能制造技术)作为重点发展的项目之一,表明我国政府有关部门对智能制造技术在制造业领域中的新内容的充分重视。预计在不远的将来,我国的智能制造技术必将有很大的进展,成为我国工业腾飞的翅膀。

第 2 章 专 家 系 统

2.1 专家系统简述

专家系统是一种在某一领域中起到人类专家作用的知识工程软件。也就是说,一个专家系统包含了有关某个特殊问题领域的知识,并且具有应用这些知识去解决该领域问题的能力。

专家系统要解决的问题一般难于形式化和转换成计算问题,这些问题往往需要由具备某一领域专门知识的专家来解决。因为专家不但掌握了处理问题的详细的专门知识,而且还具备了处理问题的经验。为了像专家那样解决问题,专家系统就要力求收集足够的专门知识。专家系统解决问题的能力主要依赖于大量的知识,所以又被称为基于知识的系统或知识基系统。

和人类专家相比,专家系统有其独到的优点,因为专家系统的知识来之于许多专家的经验,把这些经验经过整理、挑选和改造形成知识库,所以一个好的专家系统往往比单个专家具有更多的知识。专家系统不像人类专家那样要睡觉、要度假、会生病等。培养一个人类专家往往需要几十年的时间,花费大量的费用,人类专家数量有限,而专家系统可以建立很多。此外,人类专家所具有的知识会因死亡而失传,而计算机专家系统的知识却可通过版本的更新而提高。

专家系统是从 20 世纪 60 年代中期开始研制的,在 70 年代初已开发成功一些早期的专家系统,其中较著名的是斯坦福(Stanford)大学开发的专家系统 MYCIN,它主要用来处理有关感染性血液病的诊断与治疗等问题。从此以后,大量的专家系统被研制成功并在各自的领域获得应用。目前专家系统主要用于以下几个方面:

(1) 诊断 诊断系统不但可用于诊断各种疾病,在工业上还广泛应用于事故诊断,如飞机、柴油机和汽轮机的故障诊断以至于一些复杂工业系统的故障诊断。

(2) 预测 预测系统如天气预报系统,它可根据大气压、风向、风速和温度变化等因素来估计未来几天的天气趋势。又如监测报警系统,它可根据工业系统中的各种参数的变化来判断系统的运行是否正常,如果系统发生异常情况,专家系统将及时发出预警信号,以便及时加以处理。

(3) 咨询 用于某些领域的解释和咨询。常见的咨询系统有法律咨询系统,如用于回答有关雇员权利中的劳动保护、解雇、退休和最低工资等问题的法律咨询系统。

(4) 设计与制造 主要应用于智能化的计算机辅助设计(ICAD)、智能化工艺设计(CAPP)和智能计算机辅助制造(ICAM)、智能装配等领域。

(5) 管理 专家系统已成功地应用于工商管理的很多领域,如帮助银行经理决定是否向某一客户发放贷款或同意抵押的专家系统。为保险公司设计的,用于估计在某一保险中所要承担的风险系数的专家系统等。

(6) 控制 主要用于各种生产系统的生产过程,如石油、化工生产过程的智能化控制以及柔性制造系统中的单元控制器的智能控制等。

(7) 规划 专家系统现已用于许多规划系统,智能化经济规划系统、智能化建筑规划系统、资源利用规划系统等。

(8) 智能化机器教学(ICAI)。

综上所述,我们可以看到,专家系统已经在现代科学技术的很多领域发挥了重要作用。随着人工智能理论和计算机科学的飞速发展,我们有理由相信,专家系统必将在人类生活的各个领域发挥越来越重要的作用。

2 2 知识的表达与获取

知识是专家系统的核心,知识的质与量是决定一个专家系统性能好坏的主要因素。那么,什么是知识呢?专家系统的创始人 Feigenbaum 给知识下的定义为:知识是经过整理、解释、选择和转换的信息。在下面的几节中,我们将着重介绍一下知识的分类、知识的表示和知识的获取等几个方面的问题。

2 2 1 知识的分类

知识一般可分为以下几类:

(1) 叙述性知识 叙述性知识也称为事实知识,如“滚齿机属于金属切削机床”;“C6140 车床的最大加工直径为 400 毫米”;“电动机是一种把电能转换为机械能的装置”等等。

(2) 问题求解知识 问题求解知识也称为规则或领域知识,它可用产生式规则表达为:

如果.....那么

例如:如果零件表面与钻头不垂直,那么钻头会产生偏斜。

(3) 控制知识 控制知识主要描述问题求解的策略,控制问题求解的过程。

(4) 元知识 元知识是关于知识的知识。如知识的来源、范围、置信程度以及某些事实的相对重要程度等。

2 2 2 专家系统的知识表示

知识表示技术产生于人类的进化过程,它是人类文明进步的一个重要标志。知识表示的重要性可通过数学的发展史来说明。原始人只能用刻痕或绳结来计数,直到后来人们发现了计数符号,才使数学大大向前推进了一步。初期的计算机是为了数值计算而研制的,后来,各种非数值信息(例如字符串、图象和声音等)在计算机中表示的成功促进了信息处理科学的蓬勃发展,使计算机的应用扩大到社会生活的各个领域,而不像过去那样仅限于数值计算的范围内。由此可见,“知识的计算机表示”是非常重要的。事实上,如果一个问题在计算机上无法表示,就谈不上用计算机去处理它。

人工智能与计算机技术的结合产生了“知识表示”的新课题,即要求计算机模拟人脑的功能来解决一些“智能化”的问题,专家系统正是为了实现这一目标的一个机器系统。在专家系统中,计算机要处理的不再是一般的数据和函数等形式的信息,而是表示在机器中的知识。与数据处理一样,为了对知识进行处理,首先必须要解决如何表示知识,即要研究如何把人类自己的知识首先逻辑地表示出来,并最终物理地表示和存储到计算机中去,它与数的表示在数据处理中的重要地位一样,是专家系统的最重要问题之一。

一种好的知识表示方法应满足下列几点要求:

- (1) 表达能力 表达能力是指该知识表示方法能正确地、有效地将问题求解所需的各类知识表示出来;
- (2) 可访问性 可访问性是指专家系统能有效地利用知识库的知识;
- (3) 可扩充性 可扩充性指的是由该知识表示方法组成的知识库可以很方便地加以扩充;
- (4) 相容性 知识库中的知识应保持一致;
- (5) 可理解性 所表示的知识应易读,易懂,便于知识获取、知识库的检查、修改及维护;
- (6) 简洁性 知识表示应简单、明了、便于操作。

在专家系统中,采用的知识表达方法有很多种,本章将主要介绍知识的逻辑表示法、产生式规则表示法、框架表示法和面向对象的表示法。

1 . 知识的逻辑表示法

谓词逻辑是一种以符号逻辑为基础的形式语言,它接近自然语言,无二义性,并且易于在计算机内表示和推理,Prolog 语言编程则更为简单。

在谓词逻辑中要表达的知识为:事实和事实之间的关系,事实就是在某相关领域中的真实情况或事件。例如 Lathe has a spindle 表示车床有一根主轴这样的事实,这个事实是用自然语言表示的。同样这个事实可用谓词逻辑方法表示,如

has- spindle(lathe) (2 .1)

或

has(lathe spindle) (2 .2)

式(2 .1) 为一阶谓词逻辑,式(2 .2) 为二阶谓词逻辑,一般来说,有

$P(A_1, A_2, \dots, A_n)$

称为 n 阶谓词逻辑,其中 P 称为谓词, A_i 称为项。

谓词逻辑与自然语言有一种很自然的转换关系,对于二阶谓词

$P(A_1, A_2)$ 谓词逻辑

对应地有

$A_1 P A_2$ 自然语言

这种转换几乎不需要附加成分。因此把一个只有一个单词的主语、谓语和宾语的简单句转换为二阶谓词是确定的, A_1 对应主语, P 对应谓语, A_2 对应宾语。按照这种关系,

可以很方便地实现谓词逻辑和自然语言之间的转换。二阶谓词是最常用的知识表达手段。

2 . 知识的产生式规则表示法

基于规则的产生式系统是目前专家系统中最普遍的一种知识表示方法,比较成功的专家系统大多采用了这种表示模式。它类似于人们在日常生活中用到的基本判断模式:“ 如果……那么…… ”

产生式系统主要由综合数据库(global database)、产生式规则(production rules)和控制系统三部分组成。

每条规则是一个“ 条件-行动 ”产生式,且各规则之间相互作用不大。规则可用如下形式表示:

IF (触发事实 1 是真)
 (触发事实 2 是真)
 ...
 (触发事实 n 是真)
THEN (结论事实 1)
 (结论事实 2)
 ...
 (结论事实 n)

例如在某计算机辅助工艺设计专家系统中的一条规则为:

如果 表面形状是外圆柱面或外圆锥面
 加工精度等级 IT 值大于或等于 6
 表面粗糙度介于 6.3 和 40 之间
那么 可采用车削加工

产生式规则的优点是:知识以孤立的事实和独立的规则集表示,结构简单,易于模块化。规则以前提-结论的形式表述,因果关系清晰,可反映专家的求解思路,还可用来表示不确定性知识。

目前,产生式系统与初期的系统相比,已有了很大的发展,已广泛地应用于基于知识的系统中,这种系统有时也称为基于规则的系统。

3 . 知识的框架表示法

知识的框架(frame)表示法是在 70 年代初期由美国麻省理工学院 Minsky 提出的一种知识表示的方法。

知识的框架表示法可以看成是给一个或一类对象提供结构化表示的特性表的推广。每个框架有一个框架名,表示要描述的内容,在框架下面设有若干个槽。和框架一样,每个槽都有一个槽名,槽名下面有对应的取值,称为槽值,即表示特征的值。对于一些比较复杂的框架,槽下面还可以进一步分为侧面,每个侧面又可以有各自的取值。槽或侧面的

取值可以是实数值,可以由文字或其它形式的定义域,还可以是一组子程序,称为框架的程序附件,这样就形成了一层层嵌套的联接表,框架的一般形式如下:

《框架名》		
《槽名 1》	《侧面名 11》	(值 111, 值 112, ...)
	《侧面名 12》	(值 121, 值 122, ...)
	...	
...	...	
	《槽名 2》	《侧面名 21》 (值 211, 值 212, ...)
	《侧面名 22》	(值 221, 值 222, ...)
...	...	
	...	

下面是车床这一概念的框架。

车 床 架	
类型	轴数(单轴、多轴等)
	自动化程度(全自动、半自动、手动等)
	最大回转直径(320mm, 400mm, 600mm 等)
结构	床头箱(床头箱框架)
	电器控制柜 (电器控制柜框架)
	床身(床身框架)
	刀架(刀架框架)
	尾架(尾架框架)
	...
...	...
	...
床 头 框 架	
主轴部件	主轴(轴端, 轴颈, 轴尾, ...)
	轴承(前轴承, 后轴承, 中间轴承, ...)
	油封(挡油环, 油毛毡, 隔圈, ...)
箱体部件	壳体(箱盖, 箱体, 操纵面板, 标牌, ...)
	联结件(铰链, 螺钉, 螺母, ...)
传动轴部件	传动轴(第一传动轴, 第二传动轴, ...)
	齿轮(交换齿轮, 滑动齿轮, 惰轮, ...)
操纵部件	手柄(变速手柄 1, 变速手柄 2, ...)
	按钮(启动按钮, 停车按钮, ...)
润滑部件	油泵(泵体, 油泵齿轮, 滚针轴承, ...)
	油泵电机(定子, 转子, 接线柱, 冷却风扇, ...)

油管(进油管,回油管,管接头,...)

...

...

从上面的例子可以看出一个框架中可以包括其子框架,如床头箱框架、电器控制柜框架、床身框架等。框架表示法在表示知识的过程中自然形成了一个层次,使对知识的描述既可以很概括、抽象,又可以很具体、详细。

4. 知识的面向对象的表示法

面向对象的程序设计(object oriented programming)是在吸收、借鉴软件工程领域近年来新概念和新方法的基础上发展起来的一种程序设计方法,Smalltalk 和 C++ 等语言都是面向对象的程序设计语言。

面向对象的方法学认为世界由各种“对象”组成,复杂的对象可由相对比较简单对象以某种方法组成。甚至整个世界也可以从一些最原始的对象开始,经过层层组合而成。从这个意义上讲,整个世界可认为是一个复杂的对象。

所有对象被分为各种对象类,每种对象类都定义了一组“方法”(method),它们实际上可看成允许作用于该类对象上的各种操作,对该类中的对象操作都可应用相应的“方法”于该对象来实现。这种操作在面向对象方法学中被称为“送一个消息(message)给某对象”。

对象的特性之一是其封装性,对象的外部接口使外界只能知道对象的外部特征,具有哪些处理能力;对象的内部状态及处理的具体实现对外部是不可见的,即所谓的信息隐藏。

一个类可以有超类,即比它更抽象、更一般的类;也可以有子类,即比它更具体、更特殊的类,由此可形成一个层次结构。类的层次结构的另一个重要特性就是继承性,即一个类可以继承在层次结构中位于其上层的所有祖先超类的全部属性与方法。

在面向对象表示中,类和类继承是十分重要的概念,类由数据和方法组成,它描述了一组具有相同属性和操作的对象。每一个对象都属于某一类,每个对象都可由相关类生成,类生成对象的过程称为实例化。一个类拥有另一个类的全部变量与操作,这种拥有就是继承,继承是面向对象表示法的主要推理形式。

2.2.3 知识获取

知识获取是一个向专家寻求知识的过程。

作为某一领域的专家,一般要花很多年的时间甚至毕生的精力来获得和积累他所从事专业领域的知识。要从专家那里获取这些信息需要大量的时间和精力,特别是在专家们一时无法表达他们是如何作出某种决定时。因为有时专家们是凭直觉去这样做的。例如,一个车工技师可能凭感觉就可确定车刀是否已磨损,但却一时无法解释他如何确定刀具已磨损的。通过进一步交谈和分析,车工技师可能会较详细地描述刀具前刀面、后刀面和刀尖的磨损症状,铁屑颜色,切削声音的变化,机床的振动情况和零件上已加工表面

的粗糙度的变化等特征。车工技师是通过以下“趋势分析”来判断刀具是否需要更换的:这把车刀已使用了很长时间了.....工件的表面粗糙度已开始变坏.....但车刀刀尖还没有磨损.....也许可以稍微降低一点切削速度.....好,现在加工已保持稳定.....可以维持现状,直到这一刀走完,然后必须换刀。

在咨询专家时,知识工程师(通常为一个计算机软件专家)要把收集到的信息加以提取和改造,使之成为适合专家系统的概念和规则。如上述例子可归纳成以下规则:如果(IF)铁屑颜色变为蓝黑色,机床振动加剧,切削声变尖且工件的表面粗糙度已开始变坏。那么(THEN)刀具已无法继续使用。

知识工程师归纳的规则可能不很完善,因此必须用已知情况加以测试。例如,添加一些新的数据:某一特定的铁屑颜色、机床的噪声声级、零件表面反射光线的强度等。如果这一专家系统可以根据已有的知识和专家(车工技师)一样准确判断刀具的磨损状况,可以说该系统已通过这一测试。如果测试失败,知识工程师还要回到专家处,从新查找理解过程中的错误。

知识获取是人工智能研究中最重要也是最困难的问题。在专家系统的开发过程中人们发现,知识获取是最难解决的一道工序,这主要因为,目前的专家系统主要依靠知识工程师用人工方法向专家获取专门知识,为了设计一个专家系统,知识工程师必须要学习和熟悉该领域的知识。在领域专家的指导下,查询大量的文献资料,从中获取与建立知识库有关的共性知识。同时还要与领域专家密切配合,获取专家的个性知识,即启发式知识。在与专家会谈时,知识工程师要对领域知识的主要特征加以归纳并且从中抽取解决领域问题的专门知识。在这一过程中,最困难的问题就在于在很多情况下,领域专家往往自己也很难把自己的知识讲清楚,专家们在很多场合下知道如何去解决面对的问题,但却说不出为什么要这样做,这就需要知识工程师与领域专家密切配合,反复讨论,从中获取有用的知识。另一个知识获取的困难在于领域专家们的知识可能相互矛盾,在这种情况下,就有必要召集专家们共同讨论,以取得一致性的意见。

2 3 推理机

人类解决问题的能力主要体现在两个方面:一是人类拥有大量的知识;二是人类具有选择知识和应用知识的能力。这种选择知识和应用知识的过程称为推理。

基于知识的推理以已有知识的利用为前提。在基于专门领域知识的专家系统推理中,推理所采用的方法在一定程度上取决于所采用的知识表示方法。根据知识表示方法的特点,我们可以把基于知识的推理分为基于规则的推理、基于框架的推理、基于谓词逻辑的推理、基于语义网络的推理等。

评价推理机的一般标准是推理的效果和推理的效率。推理的效果是指选择和应用知识的准确性;推理的效率是指问题求解所需的时间。

下面简单介绍一下正向推理控制策略、逆向推理控制策略。

2 3 1 正向推理控制策略

正向推理又称正向链控制,它从已知事实出发,逐步逼近结论。首先,我们举一个正向推理的例子:设想我们在加工一个零件时,在切削过程中发现铁屑颜色变深,这种情况在加工过程中经常发生,特别是当刀具使用时间比较长和工件材料比较难于加工时。现在,我们把这个问题用比较一般的方式再叙述一遍:我们面临一种状态(铁屑颜色变深),并想知道这种状态将导致的后果。首先,我们提出现在的状态是铁屑颜色变深,然后,我们思考一下所有与这一现象有关的所有规则,如:

规则 1

如果铁屑颜色变深,那么刀具即将磨损。

规则 2

如果刀具即将磨损,那么必须停车并更换刀片。

让我们来看一下我们是怎样从这一现象得出必须停车并更换刀片这样一个结论的。在叙述这一系列事件之时,要始终把“正向”和“链”两个词放在心中。首先触发这一系列事件的条件是铁屑颜色变深,即把这一事实与规则 1 中的如果(IF)部分相匹配,从而得出刀具即将磨损的结论。现在推理链继续向前推进,刀具将会磨损这一结论将作为一个新的条件与规则 2 中的如果(IF)部分相匹配,从而得出必须停车并更换刀片的结论。

从上例中我们可以看到,上述推理过程是从条件(IF)出发,然后推出结论(THEN),在正向推理过程中,一般从初始条件开始,系统扫描初始条件,并与最低层次的结论匹配,并把匹配成功的结论事实作为一个新的条件,继续向前推理,直到达到最后的目标。

2 3 2 逆向推理控制策略

逆向推理又叫反向链控制。它的基本思想是从欲证明的目标出发,检查每条规则的结论部分是否含有所需要的目标,建立为证实目标所需要的一些事实。如果在知识库中没有这样的事实,则建立一些子目标,这个过程一直持续到找出所需的事实为止,此时便已证明原始目标或者原始目标为假的结论。

这里,我们也举一个例子来说明问题:如果我们使用的车床主轴突然无法运转,其原因是什么?是因为车间停电还是因为保险丝熔断?在这里,我们已经有了结论,即车床主轴无法转动,现在需要的是去寻找故障的原因。车床主轴无法转动的原因可能很多,如车间停电、保险丝熔断、主电机故障、电磁离合器损坏、启动按钮接触不良等等。把常见的原因总结成规则,如

规则 1

如果车间停电,那么车床主电机将无法启动。

规则 2

如果三相保险丝中有一相或两相熔断,那么车床主电机将无法启动。

...

规则 n

如果主电机无法启动,那么车床主轴将无法转动。

维修工在维修这台车床时,如果采用反向推理的原则,他可以首先假定车间停电,如果经检查,车间电源没有问题,然后他再去检查保险丝,如果发现保险丝中有一相或两相熔断,根据规则 2,维修工可以推断出电机无法启动,然后根据规则 n ,可推断出车床主轴无法转动的结论。如果经检查,保险丝没有问题,那么维修人员可以再假设其它方面出了故障,通过检查来确定假设是否成立。这样一直到找出一个结论或确定无解为止。从上例中可以看到,逆向推理与正向推理的不同点之一就是逆向推理要不断建立子目标,并为这些子目标寻找解决答案,从而解释顶层最终目标。

2.4 专家系统开发工具

如果决定在某一领域内开发一个用于帮助或取代人类专家的知识系统,在建立专家系统之前,必须首先考虑所采用的软件。

一般来说,目前用来开发专家系统的软件有以下五类:通用计算机高级语言、人工智能语言、知识工程语言、专家系统外壳和专家系统开发环境。

计算机语言的种类很多,目前常用的计算机高级语言如 BASIC, FORTRAN, PASCAL, C, MODULA-2 和 Ada 等。这些高级语言一般都用来开发专家系统,实际上,目前有许多著名的专家系统就是用 C 语言开发成功的。由于近年来面向对象的程序设计(object oriented programming)技术的发展,一些面向对象的设计语言,如 SmallTalk, C++ 等已在开发面向对象的专家系统工具中发挥了十分重要的作用。

人工智能语言是各种人工智能机器或知识工程系统的重要工具。利用人工智能语言,采用知识表达、知识推理、知识获取等技术与方法,设计和编写相应的程序,是构成各种人工智能系统,实现人工智能的应用的主要方法之一。目前常用的人工智能语言主要有 Prolog 和 LISP 等。

知识工程语言是专为开发专家系统及其它以知识为基础的系统而设计、研制的语言。它们一般以一种或两种比较成熟的知识表示方式为基础,如框架、产生式规则等,因而用知识工程语言书写知识库比较方便。这类语言的实现系统一般附有对相应的知识库进行推理的推理机。知识工程语言有时也称为专家系统工具,如 M.I, VP-XPRT, CLIPS 和 GEST 等。这类语言的缺点是难以实现精细的符号编程技巧。

专家系统外壳是利用已经开发成功,并在实际应用中被证明确实有效的专家系统,抽去它的具体知识内容,保留它的基本系统结构和功能。如著名的诊断专家系统 MYCIN 开发者完成了该系统的开发以后,他们认识到这个系统有两个完全不同的部分:知识库和推理机。知识库中包含了医疗诊断方面的知识,而推理机则是一个通用的反向链规则鉴别器。基于下述设想,即如果把专家系统中的知识库中关于医疗诊断方面的知识抽掉,换上其它知识,如某种柴油机的故障知识,那么,就可以很快建立一个新的专家系统。把这种抽掉知识库的专家系统进一步进行改造,就形成了一个专家系统外壳。

EMYCIN 就是一个用来建立类似 MYCIN 知识结构的咨询系统的专家系统外壳。EMYCIN 要求用来表示知识的对象、属性值和规则排列成关系表的形式,如同软件系统

要求其处理的数据具有一定的格式一样。EMYCIN 含有知识库推理及引导用户咨询所需要的所有构件,包括编辑器和排错手段,以帮助知识工程师建立专家系统。EMYCIN 是一种专家系统外壳,其应用范围主要局限于诊断。类似的专家系统外壳还有 KAS, AGE 等。专家系统外壳的主要缺点是专家系统外壳往往带来原来应用领域的局限性,其中有一些很难适合其它领域的需要。

专家系统开发环境是一个比较新的概念,一个好的专家系统开发环境应包括以下几方面内容:

- 应能提供比较多的知识表达方式,如规则、框架、逻辑、对象等;
- 不精确推理的多样性;
- 知识获取手段的多样性;
- 辅助工具的多样性,包括数据库访问、电子表格、作图等功能;
- 友好的用户界面,包括调试功能、解释功能、图形和自然语言接口等;
- 与应用领域有关的多样性。

在国外已知的专家系统开发工具中,比较靠近专家系统开发环境的有 GURU, Knowledge, ProKappa, Craft 等。国内由中国科学院数学研究所、浙江大学、武汉大学和原机电部 15 所等联合开发的“天马”也属于这类专家系统开发环境。

由于本书篇幅有限,我们在这里只对人工智能语言 Prolog 和知识工程语言 GEST 作一简单介绍。

2 5 Prolog 语言简介

Prolog (programming in logic) 语言是一种编写和探索人工智能应用程序的优秀语言,它和目前所用的大多数过程式的语言不同,Prolog 语言是一种逻辑型语言,Prolog 语言于 70 年代初期在欧洲被提出,以后日本政府又宣布 Prolog 语言作为第五代计算机的核心语言,目前微机上常用的 Prolog 语言主要有 Turbo Prolog 和 TPC 等几种。

Prolog 是以一阶谓词逻辑的 Horn 子集为基本语法,以 Robinson 的归结原理为语义,加上深度优先的控制策略而形成的一种通用人工智能语言。Prolog 在知识表达方面有明显优势,它比 LISP 语言更具有描述性语言的特点;它内设的模式匹配和回溯功能使编程人员的程序编制变得十分简洁。用这种语言书写的程序描述了什么条件将会产生什么结果,把程序员从对每一个问题找出一种解法的负担中解脱出来。和一些过程式语言相比较,在解决一个问题时,它们往往只需要很少代码,在过程式语言编写的程序里,大部分代码的作用是控制程序执行的步骤和对数据的操作,然而在 Prolog 中,程序流大多隐含在语言内部或者由程序逻辑说明。

下面,我们将简单介绍一下这种语言。

2 5 1 Turbo Prolog 的程序结构

一个完整的 Turbo Prolog 程序包括常量段、域段、数据库段、谓词段、目标段和子句段。各段以其相应的关键字 constants, domains, database, predicates, goal 和 clauses 来

标识。在程序的首部可以设置指示编译程序执行特定任务的编译指令,此外,可在程序中设置注解,以提高程序的可读性。

按照这些段在程序中出现的次序,一个 Turbo Prolog 程序的完整结构如下:

```
/ * ----- */
/ *           注           释           */
/ * ----- */
```

编译指令...

constants
 (常量说明)

domains
 (域段)

database
 (数据库说明)

predicates
 (谓词语句)

goal
 (目标语句)

clauses
 (子句语句)

各部分的功能如下:

(1) 常量段 在 Turbo Prolog 程序中可以说明和使用符号常量,常量说明段由关键字 constants 指出,其后即是说明本身。例如

```
constants
zero = 0
one = 1
pi = 3 .141592653
yello = 2
```

对程序进行编译之前,Turbo Prolog 将把每个常量换为与之对应的实际串。

(2) 域段 域段包括域说明,用于区分不同类型的数据。在 Prolog 系统中,关系的对象(谓词的变元)属于域,它可以是标准域,或者被指定的特殊域。

域段有两个用处:第一,可以对域给出有意义的名字,即使它们在内部的域是相同的;第二,特殊域说明可用于说明标准域中没有定义的数据结构。

当需要分清谓词段的各个部分时,有时域说明就比较有用。说明自己的域有助于通过给变元类型取一个名字来对谓词作文档。如:要表示一块规格为 H031030 的组合夹具基础板,其外形尺寸用其长、宽、高来表示。如果用标准域,可以用下面的谓词来说明:

```
base - plate(integer, integer, integer)
```

这个谓词说明是正确的。但如果其他编程人员要对该程序进行维护,上述谓词的说明就不能说是很清楚了,而下面的谓词说明就可以很清楚地表达谓词中的每个变元。

```
domains
```

```
length, width, hight = integer
```

```
predicats
```

```
base - plate(length, width, hight)
```

(3) 数据库段 数据库段包含一些数据库语句,它们是由于动态数据库管理的谓词,如果程序不需要数据库,该部分可省略。

(4) 谓词段 如果要在 Prolog 程序中定义自己的谓词,必须将该谓词在谓词中说明,在说明时,要告诉 Prolog 该谓词变元的域。

Prolog 本身具有丰富的内部谓词,以 Turbo Prolog 为例,如谓词 write, makewindow, readln 和许多图形功能方面的谓词,在 Turbo Prolog 中使用任何内部谓词都不必说明。

事实和规则定义谓词,事实的谓词段只是列出每个谓词名,说明谓词变元的类型(域)。子句段是程序的核心,由于对事实和规则中的对象(变元)作了类型说明, Turbo Prolog 的效率大大提高。

(5) 目标段 程序的目标部分说明程序的目的。目标段基本上和规则中的体是一样的,只是一个子目标表。两者不同的地方是目标表关键字不许跟“:-”,当程序运行时, Prolog 自动执行目标。

(6) 子句段 子句段是 Prolog 程序的核心,在子句段中,可以写入事实和规则。 Turbo Prolog 对这些事实和规则进行操作,试图找出满足条件的目标。为了满足一个目标, Prolog 从子句段的顶部开始,查看每一个事实以搜索一个匹配。当 Turbo Prolog 从上至下搜索子句段时,对于每个匹配当前子目标的子句都作内部指针标记。如果该子句不在导出问题解的逻辑路径中的话, Prolog 返回到设置指针的地方,再搜索另一个匹配,这就是我们以后要说的回溯。

2 5 2 Prolog 中的基本编程元素和语法

Prolog 语言的最基本成分是一个事实,它对有关的某个目标集以及它们之间的关系作出一个陈述。

1. 事实

形式逻辑谓词用来描述事实,在 Prolog 语法中谓词符号和常量符号必须以小写字母开始,而变量则必须以大写字母开始。

在 Prolog 中,事实“lathe has spindles”可以表示为

```
has(lathe, spindles)
```

这是一个简单事实,它表明,车床有主轴。还可以认为这一事实描述了车床与主轴之

间的关系,这一关系为“ has”。

又如事实“ machine- 1 is operating ”可表示为:

```
operating(machine- 1)
```

它表示 1 号机床正在运转。

```
broken(machine- 2)
```

这一事实表明了机床的目前状态,即 2 号机床目前处于故障状态。

在用 Prolog 描写事实时,每一事实都应含有一个关系和一个或多个参量。首先写关系,然后将参量写在括号里。

在用 Prolog 编写专家系统程序时,采集事实是一件十分重要的工作,如我们在编写一个组合夹具拼装设计专家系统时,必须采集有关各种组合夹具元件的事实,例如某种孔系组合夹具由基础件、支撑件、定位件、压紧件等组成,对于组合夹具中的元件事实,可采用谓词逻辑以如下形式来描述:

元件名(元件代码,参量 1,参量 2,...)

以基础件中的长方形基础板为例,如一种代码为 code 其关键尺寸长×宽×高为 $L \times W \times H$ 的基础板可表示为:

```
base- plate(code, l, w, h)
```

其中 base- plate 是谓词,它表示参量性质或多个参量之间的关系。在事实库中,这一系列的基础板可表达为:

```
base- plate(h031030,380,360,30)
```

```
base- plate(h031031,440,360,35)
```

```
...
```

像任何形式逻辑系统一样,它有一个任意的,但对每个参量都有一个固定的解释,对于编程者来说,有必要清楚地了解这些解释,以避免数据的不一致性。

2. 规则

除了事实以外,专家系统中还必须要规则,它们描述了领域中目标的相互关系。

在 Prolog 中,规则主要由两个部分“条件”与“结论”构成,其一般形式为:

结论:-条件

其中“:-”代表“如果”(if),事实上关键词“if”在 Prolog 程序中可与符号“:-”相互替换。如规则

```
flow(water) :-open(valve)
```

可表达为:

如果把阀门打开,那么水将流出来。也可写成

```
flow(water) if open(valve)
```


例如,选择用于立式加工中心的组合夹具长方形基础板可采用如下规则:

```
select- base- plate :-  
workpiece(W- length,W- width, - , - ),  
base- plate(B- name,B- length,B- width, - ),  
B- length > W- length + 180,  
B- width > W- width + 180,
```

其含义是,如果有一长方型基础板,其长度和宽度都比工件长 180 mm 以上,则选用此基础板。

3 . 变量

Prolog 中的变量用来表示未确定的领域元素,与其它一些程序设计语言中的变量特性相比,Prolog 中的变量有一些明显不同之处。在 Prolog 中,变量只有受约束(例示)和不受约束(非例示)两种情况,在不受约束时,变量没有值,它既非零也非空,这时变量可称为非约束变量或自由变量。当变量与某一对象匹配时,变量的值就被约束为该对象的值。

在简单的询问中,可以用变量提问 Prolog ,让它找出哪一种设备有主轴。如:

```
has( X,spindles)
```

这个询问使用了字母 X 作为变量来指示一个未知的设备,变量名必须以大写字母开始,接任意数量的字母(大写或小写)、数字、下划线。例如,下面的变量都是合法的:

```
Machine- tool  
Date- 1996
```

而下面的变量是不合法的:

```
"broken"  
operation  
2-Direction
```

适当地选择变量名可提高程序的可读性,例如,如果把前例中的 has(X, spindles) 改为 has(Equipment, spindles) 其可读性就要好得多。因为 Equipment 比 X 更具有实际意义。

4 . 求解问题

假定 Prolog 程序有事实

```
is (lathe, machine- tool) (车床是一种机床)
```

和规则

```
has (Equipment, spindle):- is(Equipment, machine- tool) (如果某设备是一台机床,  
那么它应该有主轴)
```

如果在 Prolog 的目标段给出
has (lathe, What) (车床上应有何部件 ?)
Prolog 将给出答案
What = spindle (车床上应有主轴)

2 5 3 Prolog 的工作方式

Prolog 的工作过程实际上是一个机器逆向推理的过程。当用户向机器提问时,系统顺序搜索数据库,逐项寻找与目标相匹配的事实或规则。如果匹配成功,所匹配的如果是事实则得到问题的答案,如果是规则,则把规则的条件部分作为子目标,反复搜索、匹配,直至找到答案,或完全失败从而问题无解为止。下面以上一个例子来说明其工作细节:

在上例中,首先把目标

has (lathe, What)

和事实

is (lathe, machine- tool)

相匹配,显然,目标中的谓词项 has 和事实中的谓词不相匹配,目标与事实的匹配失败。

然后 Prolog 搜索程序的下一个项发现了规则:

has (Equipment, spindle) :- is (Equipment, machine- tool)

规则的第一部分,has (Equipment, spindle),称为规则的首部,测试的方法和目标与事实匹配的方法一致。假如 Prolog 发觉可以匹配(包括谓词和所有对象),它将进一步测试跟在 if 条件后的规则部分(这一部分称为规则体)的匹配情况。

现在 Prolog 测试目标 has (lathe, What)

和规则的首部 has (Equipment, spindle)

在匹配过程中,变量 Equipment 被约束为 lathe,变量 What 被约束为 spindle 目标与规则首部匹配成功。

下面 Prolog 要测试条件 is (Equipment, machine- tool)是否满足。因为变量现在有值 lathe,Prolog 现在必须证明或推翻条件 lathe is machine- tool 以便证明首部 lathe has spindle 是否为真,下面的问题就是考查车床是否是一种机床这样一个子目标,它是由 Prolog 内部产生的,作为目标要求解决问题的一个步骤。

子目标现在是

is (lathe, machine- tool)

检测程序中的事实与规则,Prolog 发现事实

is (lathe, machine- tool)

与子目标完全匹配,于是子目标 is (lathe, machine- tool) 成功,规则首部 has (lathe, spindle) 是确实的。这样,就成功地推出了车床有主轴这样一个结论。

2 5 4 用 Turbo Prolog 构造一个齿轮设计专家系统

齿轮传动设计的过程需要大量的专家知识,在齿轮设计中采用专家系统可用来帮助用户选择传动类型、齿轮材料、精度等级、粗糙度等级、润滑油等级以及齿轮的结构形式等。很明显,采用专家系统进行设计将有益于减轻设计人员的劳动强度,提高设计质量。

GDDES 是华东船舶工业学院自行设计的一个齿轮传动设计专家系统,该系统采用了中西文 Turbo Prolog 语言编写的一个小型专家系统。该专家系统具有友好的人机接口,以菜单方式向用户提问,在设计的过程中,可以回答设计人员的“ Why ”,“ How ”等询问,增加了系统的透明度。

1 . 齿轮传动设计专家系统的构造

GDDES 由专家系统部分和齿轮零件绘图两部分组成。专家系统部分的总体结构采用了模块化组织结构。它的设计任务由多个子任务完成,每个子任务是一个模块。在本系统中主要有以下几个主要模块:

- 传动类型选择模块;
- 主要参数选择模块;
- 材料与处理方式选择模块;
- 精度与粗糙度选择模块;
- 润滑油粘度等级选择模块;
- 主要尺寸计算模块;
- 可靠度 R 计算模块;
- 公差与偏差选择模块;
- 齿轮零件绘图模块。

在以上模块中,除齿轮零件绘图模块外,其它每一个模块都可看成是一个子专家系统,这种模块分法,有利于 Turbo Prolog 语言的模块化程序设计。GDDES 的专家系统部分的各子专家系统可对应 Turbo Prolog 语言的模块化设计。这些子模块可分别独立编辑、编译和调试运行。

GDDES 由知识库、推理机、解释子系统、黑板、调度表、知识库维护子系统、人机接口和零件绘图子系统等八大部分组成。各部分的关系如图 2 .1 所示。

2 . 知识库

在 GDDES 中采用了产生式规则的知识表达方式,如在材料选择模块中我们采用以下规则:

IF 闭式 and 软齿面 and 载荷中等冲击 and 大齿
THEN 用 合金钢 ZG35SiMn,热处理为调质,硬度为 HB225 。

在精度选择模块中有以下规则:

IF 齿轮为直齿 and 圆周速度 v 小于 15,大于 10
THEN < 齿轮精度等级为 6 级,齿面粗糙度 R_a 为 0 .4 >

图 2.1 GDDES 系统示意图

GDDES 的专家系统部分的每一个模块,都有一个子知识库。在建立子知识库时,首先要建立知识库的搜索树,以便建立规则。

在 Prolog 模块化程序设计中,材料选择子知识库和精度等级及粗糙度参数选择子知识库的规则内部表示为

材料选择:谓词名(规则号,结论,前提)

精度及粗糙度选择:谓词名([齿轮形状],[圆周速度 1][圆周速度 2]),[[精度等级],[粗糙度参数]])

其它模块的子知识库的规则表示略。每一子知识库都用一个内部动态数据库表示。可用谓词 save 和 consult 分别把动态数据库——子知识库,以磁盘文件形式保存和装入。

3. 推理机的实现

Prolog 本身的推理机制是通过内部合一子程序以及搜索和模式匹配来完成的。GDDES 用过程控制来协调各子任务的执行顺序。过程控制是通过黑板、调度表、推理机来实现的。黑板分为若干个栅区,每个栅区连接两个模块(子任务)的上下文,它是上一模块的结论记录区,又是下一模块的驱动数据区。系统规定,任一模块可以到上一栅区去提取数据,但不能到下一级栅区去提取信息。根据这个原则,用户提供的技术要求(数据)放在第一栅区,可以被任意模块提取,当在栅区中寻找不到证据,可向用户提问。调度表由一组元规则构成,推理机根据元规则协调各子任务的执行顺序。在 Prolog 中,用全局动态数据存放黑板。调度表中每条元规则皆是一张各子任务顺序的表。各子任务(模块)中的元规则用以控制各子任务的内部推理,可用局部谓词表示。其推理方式一般采用正向推理。各子任务内部推理得到的中间结果可用局部动态数据库表示。

2.6 知识工程语言 GEST

CLIPS (C Language Intergrated Production System)是美国宇航局 80 年代后期推出的一种知识工程语言,可广泛用于工业、商业和管理等各个领域。它具有简单实用、通用性强、运行速度快、集成性好等优点,其汉化版本已在我国得到一定程度的应用。知识工程语言 GEST(Graphical Expert System Tool)是北京理工大学以知识工程语言 CLIPS 为基础,把计算机图形技术和人工智能技术有机地融合在一起所形成的一种新颖的知识工程语言,以适应过程控制、生产管理等多种领域的需要。由于篇幅限制,我们在这里只能对这一知识工程语言作一简单介绍,以便读者对知识工程语言有一个较为具体的概念。

2.6.1 知识工程语言 GEST 的基本组成和语法

GEST 是用于开发基于规则的知识工程语言,其基本组成有三大部分:

事实库。用于存储数据(事实)。

知识库。用于存储规则。

推理机。用于控制推理过程,使程序运行。

在 GEST 中最重要的语法单位是“字段”(field),字段有三种类型:

(1) 字(word) 在 GEST 中,字以可打印的 ASCII 字符开始,后随零个或多个字符。但下列字符及其组合不得作为字的开始字符:

< | & \$? + - ()

此外,字中间也不能包含下列任一符号:

< | & () ;

这是因为以上 6 个符号在 GEST 中用作字的限界符 (delimiter),表示字的结束。根据以上规定,以下均为有效的 GEST 字:

machine. center

control

FMS

caution

GEST 认为同一字母的大小写是两个不同的字符。因此,在 GEST 中下面是三个不同的字:

TOOL

tool

Tool

在 GEST 中,空格一般用作字段的分界。在这种情况下,多个连续空格只当作一个空格。

(2) 串(string) 在 GEST 中,串以双引号开始和结束,双引号本身是串的一部分,其间可以有任意个字符。例如:

```
"Start the machine"
"! ? # * * * "
"< - ; ( ) + - "
```

(3) 数值(number) 在 GEST 中,所有数,包括整型数,都以浮点形式存储。数值字段包括三部分:符号、值、指数,但符号和指数这两部分是任选的,可有可无。根据以上规定,以下都是 GEST 中的有效数值字段:

```
1
1.2
.7
+3
-5
3.8e8
```

在 GEST 中还有一些用来描述命令和结构的常用符号,现简单介绍如下:

- (1) 圆括号 在 GEST 中,采用了类似 LISP 语言的风格,其所有的命令、事实表和规则等都采用圆括号,在使用时要注意括号的配对,以免出现语法错误。
- (2) 方括号 在圆括号中,用方括号括住的内容为任选项。
- (3) 尖括号 在圆括号中,用单尖括号 括住的内容为必选的,而用双尖括号括住的内容则代表任意多个(包括零个)与其内容相对应的值,三尖括号是双尖括号的一种特例,表示任意多个,但不包括零个与其内容相对应的值。
- (4) 大括号 表示必须采用其中一项输入。

2 6 2 事实库及其操作

为了建立一个专家系统,必须有据以推理的数据或信息,这在 GEST 中叫做“事实”(fact)。事实是由配对的左右括号括住的一个或多个有关联的字段组成的,例如:

```
(condition cutting-tool-is-broken)
(condition power-failure)
(condition machine-tool-is-out-of-order)
```

其中第一个字段名 condition 说明事实的性质(状态)是共同的,称为“关系名(relation name)”;第二个字段具体说明各种不同的加工状态,如刀具磨损、停电、机床故障等等。

给事实库即事实表(fact-list) 中加入事实有以下几种方法:

- (1) 直接通过 assert 命令 在 GEST 提示符下键入

```
(assert condition cutting-tool-is-broken)
(condition power-failure)
(condition machine-tool-is-out-of-order) )
```

这样,我们向事实库中同时加入了刀具磨损、停电和机床故障三个事实。

- (2) 通过 deffacts 命令结构 assert 命令一般只用于在程序中输入少量事实。对

于大量的原始事实,还可以通过 deffacts 输入事实库。例如:

```
(defact condition
  (condition cutting-tool-is-worn)
  (condition power-failure)
  (condition machine-tool-is-broken))
```

这个结构定义了“ condition ”这个 deffacts,其中包括刀具磨损、停电和机床故障等三个事实。在 deffacts 定义的事实在输入时并不直接进入事实库,而是要通过系统复位命令才进入事实库:

```
(reset)
```

reset 命令不带任何参数。它是用来启动和重新启动专家系统程序的,这一命令撤消事实表中的所有事实,也撤消系统中所有已被激活的规则。在一般情况下,它还自动生成一个初始事实 (initial-fact)。

(3) 通过 deftemplate 命令结构 除了以上两种输入事实的命令结构以外,GEST 还提供了一种类似框架结构的模板输入功能,其结构为

```
(deftemplate deftemplate-name
  《 slots 》)
```

在前面所述的输入方式中,事实所包含的信息和位置密切相关,因此,知识工程师在输入信息时,必须牢记这个信息位于事实中的位置,否则就会引起差错。采用 deftemplate 结构,可以使事实和事实中的每个字段都有名可依,在输入信息时可按其槽名输入,不会造成混淆。

在格式中的“ 槽 ”(slots) 由字段(field) 的字段名、类型和缺省值组成,其具体格式如下:

```
(field field-name
  [(type field-type ) ]
  [(default ? NONE, 《default-value》) ])
```

如我们要定义一个电机样板,该样板由电机种类、功率和重量三个字段组成,则该电机样板可如下定义:

```
(deftemplate motor
  (field kind
    (type STRING)
    (default AC-motor) )

  (field power
    (default ? NONE) )

  (field weight
    (type NUMBER)
    (default ? NONE) ) )
```

其中 (default ? NONE) 表示该字段无缺省值, 数据类型可以是字 WORD 串 STRING 或数值 STRING。

以上样板定义了电机的框架, 该框架由三个槽(字段) 组成, 有了这个框架后, 在输入电机信息时, 除了电机种类可以缺省外, 必须给出电机的功率和重量, 有了上述样板后, 可输入以下两种电机:

```
(assert motor
      (kind DC-motor)
      (power 1)
      (weight 2 3) )
(motor
  (power 1 .1)
  (weight 1 8) ) )
```

这两种电机分别为功率为 1 千瓦, 重量为 2 .3 公斤的直流电机和功率为 1 .1 千瓦, 重量为 1 .8 公斤的交流电机。

(4) 通过文件输入 为了方便地输入事实, GEST 中还可利用其它编辑软件把事实先输入一个文件, 然后再用装入命令一次装入:

```
(load-facts file-name )
```

2 6 3 规则库及其操作

为完成一定的任务, 一个专家系统必须要有其规则库, 在规则库中保存着某一领域的人类专家在实践中积累起来的经验, 在 GEST 中, 这种经验以“ IF ...THEN ...”的形式表示。

在一般的高级编程语言中, 如 BASIC, FORTRAN, PASCAL, C, Ada 等都有 IF ... THEN ... 语句。一个 IF ...THEN ...规则可用如下自然语言和计算机语言的混合来表达:

```
IF      某些条件成立
THEN  执行下述动作
```

然而, 由于图形知识工程语言 GEST 有着类似于 LISP 的句法, GEST 中不采用 IF THEN 等关键字, 尽管如此, 如果把类似 IF THEN 的词记在心里, 那么将规则由自然语言译成 GEST 规则并不难。

规则输入系统可以直接从键盘输入, 也可从原先建立 GEST 文件装入系统。首先, 我们以下述规则为例, 了解 GEST 如何键入一条机加工规则的。

```
IF cutting tool is broken then stop the machine .
```

我们可在 GEST > 的提示符下键入

```
(defrule operation “ rule1 ”
(cutting-tool-is-broken)
= >
```



```
(printout t ( stop-the-machine) ) )
```

这个规则的意思是,如果切削刀具损坏,那么停下机床。

从上例可以看到,一个规则包括以下几个部分:

整个规则必须用括号括起来;

规则必须以关键字 `defrule` 开始;

`defrule` 之后必须是规则名,名字可以是任何有效的符号元素,在本规则中,规则名可以是任何有效的字段,在这个规则中,规则名是 `operation`(操作);

在规则之后可以有一个任选的用双引号括起的注释,注释通常用来描述规则的用途或程序员希望标注的其它信息;

在注释之后是一个或多个条件元素,每个条件元素包含一个或多个字段,在本规则中,每个条件元素包含一个或多个字段;在规则 `operation` 中,条件元素为单字段元素 `cutting-tool-is-broken`。

GEST 设法根据事实库中的事实来满足规则中的条件元素。如果一个规则中所有的条件元素都和事实相匹配,那么,该规则将被激活,并被放入待议事件表(`agenda`)中。这个待议事件表是一个被激活规则的集合。GEST 执行待议事件表中的规则右部所规定的动作,这叫做点燃(`fire`)。如果待议事件表中没有规则,程序一般将停止运行。如果待议事件表中有多个规则,GEST 将按照优先级(`salience`)的大小对它们进行排队,并选择优先级最高的规则点燃。

2 6 4 模式匹配

模式匹配是推理机制中的关键,GEST 提供了许多强有力的手段,这里主要介绍一下变量和字段通配符。

(1)变量 在 GEST 中,变量的形式是问号后面跟一个表意的字段名,如

```
? name
```

```
? node
```

利用变量在规则的左面约束一个值,然后在规则右面把这个值赋给相应变量而断言事实,这是变量的一个最普通的用途。例如

```
(de rule fish
  (is-a-fish ? fish-name)
  = >
  (assert(live-in-water ? fish-name)))
```

如果断言事实(`is-a-fish shark`)且运行该程序,则该规则将断言事实(`live-in-water shark`)因为变量 `? fish-name` 已约束到 `shark`。

变量还常用于屏幕输出(`printout`),例如

```
(d frule fish
  (is-a-fish ? fish-name)
  = >
```

```
(printout t( ? fish-name "lives in water ." ))
```

如果断言事实(is-a-fish Shark)且执行该规则,屏幕上将显示:

```
Shark lives in water .
```

变量的另一个非常重要的用途,是把变量同一个事实地址联系起来,事实地址用左箭头操作符“<-”来指定。

举一个从规则中删除事件的例子,打入下例:

```
(de rule chang-number
  ? old-fact <-(tel-number 021-422290)
  =>
  (retract ? old-fact)
  (assert(tel-number 021-8422290)))
```

这时如果断言事实(tel-number 021-422290) 并执行规则,则旧事实(tel-number 021-422290)将从事实表中撤消,而新事实(tel-number 021-8422290) 将被加入。

(2) 通配符 通配符有两种,单字段通配符和多字段通配符,单字段通配符用 ? 来表示,而多字段通配符则用 \$? 来表示。

例如下列规则:

```
(defrule firstname
  (John $ ?)
  => (printout t "There is a person ,his first name is John" crlf))
```

如果输入以下事实

```
(assert(John smith))
(assert(Tom ))
(assert(John henry))
```

当打入运行时,规则将点燃 2 次。

2 6 5 待议事件表及其执行

在专家系统运行时,要把规则中的各个条件同事实库中的事实进行匹配,如果事实库中的事实能与某一规则左部的所有条件项匹配,则该条规则将被激活(activate)并被置于待议事件表(agenda)中。在调试时可以用(agenda)命令(watch activations)来检查和监视待议事件表中的规则是否被点燃(fire)。

如果待议事件表中有多项规则,GEST 将按优先级(salience)的大小对它们进行排队,并选择优先级最高的规则点燃。

GEST 程序的运行是通过 run 命令起动的:

```
(run [ limit ])
```

其中任选整数参数 limit 是用来限制本次起动中点燃规则的极限数量。如果 limit 为

- 1 或省略这一任选项,则待议事件表中的规则都将被点燃,直到待议事件表变空为止。

2 6 6 一个简单的应用实例

有了事实库和规则库,我们可以着手编写一个应用实例,这个小型专家系统主要用来说明如何用专家系统语言来实现机械零件的计算机辅助成组编码。当然,真正的智能化计算机辅助成组编码系统要复杂得多。

成组编码系统采用多位数字码的排列组合描述零件。前面若干位用来描述零件形状和结构特征,叫做形状码;后面几位用来描述零件的尺寸、材料、毛坯形式和精度,称为辅助码。为了使问题简化,我们这里只以有两位编码的简易编码系统为例,来说明知识工程语言在成组技术中的应用。

该简易编码系统的零件分类表如下:

表 2 1 零件编码示例

第一码位		第二码位	
特征	代码	特征	代码
回转体	0	无孔	0
		通孔	1
		盲孔	1
非回转体	1	板类	3
		箱体类	4
		其它类	5

上述分类问题实质上是一种决策树,适用于用决策树来解决问题的一个主要特点是这类问题有预先可以确定的解答,如动物的分类和疾病的诊断等问题。

决策树像一棵倒长的树,它由结点和树枝组成。其顶部为根结点,结点可分为判定结点和答案结点(叶子结点),树枝表示结点之间的联系。对于上述零件的分类编码问题,我们可以用一棵二叉决策树来解决,即假定对于每个判定结点上的问题的回答只能是“是”或“否”,因此,遍历这棵树以达到一个答案的过程是十分简单的。推理过程由根结点开始,视对问题的回答是“是”或“否”逐步由左分枝或右分枝下推到子女结点直到下一个答案结点。其算法的伪码程序如下:

解二叉树算法过程。

把当前位置设置在根结点上。

如果当前位置是一个判定结点,那么进行如下工作:

根据结点的内容提问。

如果其回答为“是”,则把当前结点设置在“是”这一分枝上,否则把当前结点设置在“否”这一分枝上。

直到当前结点为答案结点。

从当前结点中得到结论。

结束。

决策树的每一个结点可用一个事实表示,因为答案结点和判定结点上信息类型不同,所以要用不同的样板,答案结点的样板为:

```
(node name answer value )
```

其中 name 是结点的唯一的名, value 表示存储在结点中的答案。answer 这个字用来表示答案结点。

判定结点的样板如下:

```
(node name decision question yes-node no-node )
```

其中 name 也是结点唯一的名, question 是经过这个结点时提出的问题, yes-node 和 no-node 分别是对问题予以肯定或否定回答时要处理的结点。

按照以上样板,我们把成组编码分类系统的事实保存在名为 code .dat 的事实库中,事实库中的内容如下:

```
(node root decision "Is the workpiece cylindrical ?" node1 node2)
```

```
(node node1 decision "Has it a hole ?" node3 node4 )
```

```
(node node3 decision "Is the hole a blind hole ?" node5 node6)
```

```
(node node2 decision "Is it a plate ?" node7 node8)
```

```
(node node8 decision "Is it a box ?" node9 node10)
```

```
(node node4 answer CODE = 00)
```

```
(node node5 answer CODE = 02)
```

```
(node node6 answer CODE = 01)
```

```
(node node7 answer CODE = 13)
```

```
(node node9 answer CODE = 14)
```

```
(node node10 answer CODE = 15)
```

其中前 4 行为判定结点,其中根结点标明 "root",其它结点都有唯一的结点名。

下面给出遍历判定树的一些规则。首先是初始化程序的规则:

```
(de rule initilize
```

```
  (not(node root $ ?))
```

```
  = >
```

```
  (load-facts "code .dat")
```

```
  (assert(current-node root)))
```

如果事实表中没有根结点,这个规则将被点燃,把事实库装入内存,并把根结点作为当前结点。

下一个规则执行判定结点所需要的动作:

```
(defrule do-decision-node
```

```
  ? node <- (current-node ? name)
```

```
  (node ? name decision ? question ? yes-branch ? no-branch)
```

```

= >
(retract ? node)
(format t "% s(yes or no)" ? question)
(bind ? answer(read))
(if(eq ? answer yes)
then(assert(current-node ? yes-branch))
else(assert(current-node ? no-branch))))

```

这个规则中的第二个模式仅当当前结点是判定结点才是匹配的。

```

(defr le do-answer-node
  ? node< -(current-node ? name)
  (node ? name answer ? value)
  = >
  (printout t "The GT code is " ? value crlf))

```

在把以上规则写入名为 "code .rul" 文件中以后,就可以运行该程序了。运行步骤和结果如下所示:

在 GEST> 提示下键入

GEST>(load " code .rul") 回车

GEST>(reset) 回车

GEST>(run) 回车

这时屏幕上会出现:

Is the workpiece cylindrical(yes or no) ?

如果我们键入 yes,屏幕上会出现:

Has it a hole ? (yes or no)

如果键入 no,屏幕上会出现:

The GT code is CODE =00

即该零件的成组编码为 00

2.7 专家系统的开发

2.7.1 开发专家系统的特点

从国外研制专家系统的经验来看,一个成功的专家系统往往离不开知识工程师和领域专家的密切合作,这是开发专家系统的主要特点之一。专家系统和一般的软件系统的不同之处在于:对于一般的应用系统而言,由于系统涉及的知识比较简单,系统开发人员可以在比较短的时间内熟悉和掌握这些知识,程序设计人员往往同时兼任上述两种角色;对于专家系统来说,情况则大不相同,要使专家系统达到甚至在某些方面超过单个专家的水平,就必须掌握领域专家处理问题时的数量专门知识。这些经验知识往往要通过长期

的积累;另一方面,专家系统是一种很复杂的计算机程序,开发一个专家系统不但需要人工智能方面的知识,而且需要一定的计算机知识和编程技巧。因此,无论是让领域专家来取代知识工程师或者是让知识工程师来兼任领域专家的工作,其效果都不理想。事实证明,只有通过两者的密切合作,共同努力,才有可能开发出理想的专家系统。

领域专家的专门知识,大多数是长期积累的经验性知识,专家们在很多场合可以运用他们所具备的知识很快解决他们所需要解决的问题,但在短时间内将这些知识很好地整理出来通常是比较困难的。通常使用的办法是通过大量的实例,研究领域专家在解决问题时使用了哪些知识,如开发一个计算机辅助工艺设计专家系统,工厂中现有的大量工艺文件就是十分好的实例。一般来说,开发一个成功的专家系统需要用较长的时间来整理知识和构造知识库。

从上面的讨论中我们可以看出,在开发和研制专家系统时,领域专家、知识工程师和大量的实例是必不可少的三大部分。除此以外,在进行专家系统设计时应遵循以下基本原则:

(1)知识库与推理机分离 这是专家系统的基本原则,只有知识库和推理机相分离,才能实现解释功能和知识获取功能。

(2)尽量使用统一的知识表示方法 这就便于对系统中的知识进行统一处理、解释和管理,从而使专家系统的实现相对简化。

(3)推理机尽量简化 把启发性的知识尽可能地独立出来,这样既便于推理机的实现,同时也便于对问题的解释。

2.7.2 开发专家系统的基本步骤

研究专家系统的最终目的是要设计出能解决实际问题的智能系统。广义的讲,专家系统也可以认为是一种信息系统。专家系统的设计也要经过系统分析、设计和实现的过程,一般来说,专家系统的设计过程可以分为以下几个基本步骤:

1. 初步论证

这一部分的工作主要包括需求分析,要了解用户对于本专家系统的需求以及国内外对这一问题的研究状况,要明确需要解决的问题,对于我们准备解决的问题或所接受的任务来讲,它属于什么性质?有哪些主要特征、要求,以及其中的关键问题是什么?有没有解决相应问题的方法和技术?并对建立专家系统的可行性进行分析。在这一阶段,要对现阶段解决该问题的方式和水平有一个比较明确的调查。确定开发系统的人力(包括知识工程师、领域专家和程序员等)、物力、财力等,一般来说,开发一个小型专家系统约需要4~8人年,开发一个中型专家系统约需要8~16人年,而开发一个大型专家系统则需要16人以上。可行性分析还应包括对研制专家系统所需要的设备条件,研究人员的水平和经验等问题进行分析,以确定本单位是否已具备必要的研究条件以及研究人员是否能胜任这一工作。最后,还应该对系统的预期效果进行分析。经过以上初步论证后,才能开展下一步的工作。

2 . 知识的获取和整理

由于知识是建造专家系统的基础,因此在一个专家系统的建造中,解决知识如何获取是一个十分重要的问题,专家系统的创始人之一的 Feigenbaum 曾经说过,知识获取是人工智能研究的中心问题中最重要。在专家系统的开发过程中,人们发现,知识获取是所有步骤中最困难的一步。知识获取可分为自动型知识获取和非自动型知识获取,目前专家系统的知识获取主要还是采用非自动知识获取的方式,即主要由领域专家和知识工程师参与的知识获取。知识工程师的主要任务是:在领域专家的指导下,翻阅大量的有关文献、资料和手册等,从中获取与本专家系统有关的知识,同时还要花费大量时间与专家配合,获取专家的个性知识,即启发式知识。根据获取的以上两方面知识,来设计建造专家系统的知识库。

知识的整理主要是指知识的概念化和形式化。

问题的概念化主要包括以下内容:

需要解决的主要问题;

问题所涉及的主要方面;

问题可细分为哪些子问题;

问题及子问题求解所需要的知识类型;

确定需要解决的问题的求解方法及所需要的概念。

在形式化工作中,需要把概念化阶段整理出来的概念、概念之间的关系以及领域专门知识等用适合计算机表示和处理的方法描述出来,并选择与问题相应的控制策略。在形式化工作中,知识工程师担任着主要任务,他的基本任务是根据应用领域的特点,选择和建立合适的系统构造技术。

3 . 建立模型系统

在实现工作中,知识工程师就可以着手建立系统的模型系统,并将形式化的概念和知识装入系统中。形式化阶段虽已明确地确定了表示概念与事实的数据结构、推理规则以及系统的控制策略,但它们并不是可执行的程序。它们之间某些不一致的东西,只能在实现阶段暴露出来。知识工程师必须在实现系统时消除形式化阶段中存在的 inconsistency,保证系统中各个部分能有效地衔接起来,并正确地运行。

在建立模型系统阶段,不必追求系统尽善尽美,尤其是不要追求知识库完美无缺,这一工作可以留到系统的扩充和改进时完成。在这一阶段,还应尽量保持模型系统的简单化,这是为了使模型系统的修改不致引起大量的附加修改工作。在建立模型系统阶段,如果能开发一些知识的获取功能和解释功能,可以帮助知识工程师把整理出来的知识加入到知识库中,简化知识库的构造和修改。简单的解释功能还可以帮助知识工程师追查出错的原因,便于发现专家系统的缺陷。

4 . 测试和完善

系统的测试主要有两部分内容,软件本身正确性的测试和系统性能的测试。

测试和完善阶段的主要目的就是要通过使系统处理大量的各种类型的实例来检测系统的性能水平是否理想或是否正确。一旦发现系统得出错误结论,就必须追查导致系统出错的原因,是知识库不完备、不一致造成的,还是实现方案不合适导致的错误。为了更正这些错误,可能会对系统进行修改,甚至作较大的变动,以使系统不断完善。

由于在前一阶段知识的获取过程中,领域专家已基本上了解了整理知识的要求,所以这一阶段的知识获取工作可由领域专家单独进行,与系统的改进扩充工作并行,以减少系统的开发时间。但是在这一段时间内领域专家和知识工程师仍应保持密切的联系,进行讨论,交流进展情况。领域专家把整理出来的知识交给知识工程师,知识工程师检查这些知识是否符合要求,必要时还需进一步划分一下知识结构。知识工程师要向领域专家通报系统的改进情况,以便于领域专家下一步整理知识,同时还要请有关人员来评价系统的性能,提供下一步的改进意见等。

这一段是反馈比较频繁的阶段,由于知识库的内容越来越丰富,可提供测试的实例越来越多,特别是实现细节之间的不一致性,实现细节本身的不充分,大多在这一阶段暴露出来。因此在这一阶段,对系统的修改工作也比较频繁,知识工程师要对系统进行不断的改进,使之能够正确地运行和解决更多的问题。

2.7.3 专家系统的发展趋势

1. 大型专家系统和小型专家系统并行发展

专家系统的规模大小从目前来看有两个主要趋势:一个是向小型知识系统发展,它可以是由用户在知识工程师的帮助或指导下建立起来,这种系统可以用来辅助问题的分析和决策,解决较小范围的实际问题,这些专家系统很注重与用户的友善。由于开发者和使用者的关系往往比较密切,所以这类系统一般处于经常性的改进之中。另一个是向大型专家系统方向发展,这些项目往往对国民经济发展起比较重大的作用。开发这样的专家系统需要领域专家和知识工程师的全力配合,采用适合的开发工具来进行,开发前必须对知识库的规模要有充分的估计,以便论证采用何种类型的软件和硬件。

2. 各种专家系统工具和开发环境得到普遍应用

在专家系统的早期开发阶段,大多数的专家系统都是用普通计算机语言如 BASIC, FORTRAN 来编写的,以后由于微机版本的人工智能语言 LISP 和 Prolog 的引进,不少小型专家系统转向用这两种语言来进行开发。但经过一段时间的应用后,人们发现,即使是人工智能语言在描述专家系统的知识表示和控制机制方面还是不够方便,于是各种各样的知识工程语言(也有称为专家系统工具)便应运而生了。前一阶段较为流行的 OPS5, VP-EXPERT, CLIPS, OPS83, GEST, M-1 等都属于这一类的例子,这些知识工程语言和专家系统外壳不同,它们的应用范围要广泛得多,这从我们上面的介绍中也可看出。随着专家系统在各方面应用的不断深入,人们发现,上述知识工程语言还有很多不足之处,如我们在开发组合夹具智能化拼装专家系统的过程中,需要有一个组合夹具的三维元件库,而现有的知识工程语言的三维造型功能还是比较差的。由于用户需求的多样化,

如有的用户需要用数据库,有的需要各种统计和建模算法,有的需要生成电子表格,还有的提出和自然语言系统打交道。为满足以上种种要求,好的专家系统开发工具应包括上述功能,这一类专家系统开发工具目前被称为专家系统开发环境。如 MDBS 公司在 80 年代推出的 GURU 就是这样一种工具,它包括了关系数据库管理、标准的 SQL 语言、统计分析、电子表格、通用文本处理、通用报表生成、远程通讯、自然语言接口开发以及专家系统外壳等。

3. 面向对象的技术逐步被采用

面向对象的程序设计近年来受到越来越多的重视,它是吸收了软件工程领域十几年来新概念和新方法而发展起来的一种很有前途的设计方法,现已应用于计算机仿真、系统设计、数据库、操作系统、分布式系统以及专家系统等各个领域。

当前在人工智能领域,对于基于规则知识系统的设计,正在从基于规则的机制向面向对象的和利用匹配规则的混合机制发展,面向对象的专家系统开发工具的研究已成为目前专家系统领域的一个热门课题。

市场上已出现了一些这样的产品,如 IBM 公司推出的 LEVEL 5 Object,作为以面向对象为基础的换代产品。Intellcorp 公司推出的 ProKappa 也是以面向对象为基础的产品。这些产品的使用效果还有待用户进一步验证。

4. 向标准化的方向发展

目前,专家系统中的知识的表达方式种类很多,即使是同一种表达方式,如产生式规则,其形式也是千变万化的。这就产生了一个知识难以共享的问题,这个问题对于小型专家系统来说,并不显得十分重要,但是对于一个大系统来说,比如一个 CIMS 系统,如果知识库的知识无法共享,那就会造成很大的麻烦。

目前,在标准化工作方面做得比较早的是 IBM 公司,它首先开展了一个称为 AD/ Cycle and Repository 的标准产品系列计划,其中规定要对规则、对象等基本知识给出标准定义。与此同时,由 Carnegic Group, DEC, Ford, TI 和 USWest 五家公司联合建立一个组织,称为 IMKA,意思是要把知识库规范化为一种可共享的知识财富。

由日本提出的国际合作计划 IMS(智能制造系统)中对于信息的标准化、规范化和通用化也十分重视,把制造智能共享问题列为其重点研究内容之一。

5. 作为专家系统的神经网络

进入 80 年代后,虽然专家系统已经在不少专门领域显示了相当出色的工作能力,在许多场合不仅达到而且超过了人类专家的工作能力,新的不同领域的专家系统像雨后春笋般地被建造出来。但专家系统技术本身还是有不少局限性的,如处理多个领域专家知识的矛盾问题,专家系统性能的“窄台阶”效应。对于专家系统领域以内的问题,专家系统能以专家的水平来处理;但对于专家系统以外的问题,系统无法自行判别其是否已超出边界,因而得出一些荒谬的结论。现有的专家系统都几乎没有什么自学能力,因此使得专家系统的维护和管理变得十分困难。

因此,是否可以由神经网络来实现专家系统的功能?从目前的研究结果来看,这一答案是比较肯定的。不仅如此,由神经网络设计成的专家系统还具有传统专家系统所不具备的许多优点,特别是基于神经网络的专家系统在知识获取、并行推理、适应性学习、联想推理、容错能力等方面都显示了一定的优越性,而这些方面恰好是传统专家系统的缺点。

当然,神经网络的专家系统也有其本身的弱点,首先,神经网络的专家系统在目前还只适合于解决一些规模比较小的问题;其次,神经网络的专家系统的性能在很大程度上受到所选择的训练数据集的限制。训练数据的正交性和完备性如果不好,就会使系统的使用性能劣化。神经网络的专家系统的最严重问题之一,就是它没有能力来解释自己的推理过程和推理依据,也不能向用户提出必要的询问。

尽管神经网络的专家系统还存在不少问题,但是我们认为,目前的探索是十分有益的,有必要进一步开展研究。

第 3 章 神经网络

3.1 概述

神经网络的研究已有 40 多年的历史了。20 世纪 40 年代与神经网络有关的神经科学研究,诸如大脑功能的生理学、心理学研究、神经元的电生理实验等都有很大的进步和可喜的成果。该时期由 McCulloch 和 Pitts 提出了 McCulloch- Pitts 神经模型(即 MP 模型)以及由 D .Hebb 提出的一种将学习功能引入神经网络系统并改变神经元之间连接强度的 Hebb 规则。尽管该模型和规则在当时是作为一种假设提出来的,但其正确性已得到了证实,而且从其假设的提出到现在一直在神经网络模型中占有一席之地,起着重要作用。

经过 20 多年的低潮以后,80 年代开始,神经网络研究又重新兴起。神经网络系统最主要的特征是大规模并行处理,信息的分布存储,连续时间的非线性动力学,高度的容错性和鲁棒性,自组织、自学习和实时处理。目前,人工神经网络的应用主要在以下三个方面:信号处理与模式识别;知识工程或专家系统;运动过程控制。这些应用实例的共同特点是:难以用算法来描述待处理的问题,存在大量的范例可供学习。在专家系统的研究中就是如此。与传统的人工智能不同,神经网络在处理专家系统中的知识表示时用一种隐式表示,在这里知识并不像在产生式系统中那样独立表示每一规则,而是将某一问题的若干知识在同一网络中表示。知识表示表现为内部和外部两种形式,面向专家、知识工程师和用户的外部形式是一些学习范例,而由外部形式转化为面向知识库和内部编码是其关键;它不是根据一般代码转换成编译程序,而是通过机器学习完成,机器学习程序可以从范例中提取有关知识,并通过权矩阵及系统参数进行学习训练,将有关知识以网络或动力系统形式表示。多层神经网络的内部可用分块邻接权矩阵和阈值矢量描述。总之,神经网络具有很强的学习能力,将在专家系统的研究中得到广泛的应用。

3.2 神经网络的基本模型

现有的神经网络模型有几十种。随着理论研究和应用的进一步展开,新的模型还会不断涌现。本节主要介绍几种常用的神经网络模型,包括感知机、BP 模型、Hopfield 模型和 Boltzmann 机模型。

3.2.1 感知机

20 世纪 50 年代末,F .Rosenblatt 设计了“感知机”,它是一种分布并行处理系统,系统由乘加运算的处理单元构成的网络组成。图 3.1 所示一个感知机模型。

输入矢量：

$$\mathbf{X} = [x_1 \quad x_2 \quad \dots \quad x_n]$$

加权系数矢量：

$$\mathbf{W} = [w_1 \quad w_2 \quad \dots \quad w_n]$$

输出： y

其数学运算可表示为：

$$\mathbf{O} = \mathbf{W} \cdot \mathbf{X}^T = \sum_{j=1}^n w_j x_j \tag{3.1}$$

$$y = f(\mathbf{O})$$

这里 $f[\cdot]$ 是一个单调非降函数。当 $f[\cdot]$ 取线性函数时,即 $f(x) = x$ 称为线性感知机,这就是 Rosenblatt 最早提出的两层感知机模型。

$f[\cdot]$ 还可以取另外两种形式：

符号函数：

$$f(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

S 型函数：

$$f(x) = \frac{1}{1 + e^{-x}}$$

图 3.2 所示为三种变换函数。

图 3.2 感知机模型的三种变换函数

下面,我们从一种最简单的情况出发,研究利用感知机解决分类问题。假定输入的观察矢量 \mathbf{X} 是二维的,即 $\mathbf{X} = [x_0 \quad x_1]$,被观察的矢量可以分为两类,即 \mathbf{C}_1 类和 \mathbf{C}_2 类,系统只有一个输出节点 y 。在图 3.3 中给出了一个简单的例子。

现在需要解决的问题是:如何通过学习得到一组数值,使得输入 x_1, x_2 能被正确分类。这里,我们只要找出直线 $L = ax_1 + bx_2$ 即可,令: $ax_1 + bx_2 > 0$, \mathbf{X} 属于 \mathbf{C}_1 类; $ax_1 + bx_2 < 0$, \mathbf{X} 属于 \mathbf{C}_2 类,也就是

$$\begin{aligned} y = \sum_{j=1}^2 w_j x_j / x \cdot c_1 &> 0 \\ y = \sum_{j=1}^2 w_j x_j / x \cdot c_2 &< 0 \end{aligned} \tag{3.2}$$

图 3.3 二维空间中线性可分类示例

这样,就可以完成对于 \mathbf{X} 的分类。对于实际应用而言, \mathbf{C} 和 \mathbf{C} 矢量一般事先不能确定, L 也无法确定。因此,各个权系数 w 也不能事先确定,只有用足够数量的已知类别的输入矢量 \mathbf{X} 样本对感知机进行训练,让感知机自动地找到那些权系数,从而确定分割线 L ,这样的分类才有实用意义。这说明感知机在实现分类功能以前,必须有足够的样本供感知机学习或训练,因而,这是一种在教师指导下的有监督学习。

需要指出,线性感知机对于某些问题不适用,例如,异或问题。因此,人们在线性感知机的基础上又提出了新的前向多层神经网络模型——BP 模型。

3.2.2 BP 模型

Rumelhart(美国)等人于1985年提出了一种 back-propagation 神经网络算法, BP 模型,它突破了两层网络的限制,加入了隐含节点,给出了一种有效的学习算法。根据学习误差的大小,从后向前修正各层之间的连接权重,使网络的学习误差达到最小。BP 网络如图 3.4 所示。

这个算法的学习过程由正向传播和反向传播组成。在正向传播过程中,输入信息从输入层经隐含层单元逐层处理,并传向输出层,每一层神经元的状态只影响下一层神经元的状态。若在输出层不能得到期望的输出,则转入反向传播,将误差信号沿原来的连接通路返回,通过修改各层神经元权值,使得误差信号最小。在理论上,BP 网络的节点作用函数只需是处处可微的,最常用的有 Sigmoid、双曲正切、正弦函数等 S 型函数,即分别为:

图 3.4 BP 网络

$$f(x) = \frac{1}{1 + e^{-x}}$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$f(x) = \sin x$$

下面具体分析一下网络的学习过程。

假定有一个 m 层的阶层型网络,把第 k 层的第 i 个节点记为 u_i^k , u_i^k 的输入是 i_i^k , 输出是 α_i^k 。输入输出函数记为 $f(\cdot)$, 即: $\alpha_i^k = f(i_i^k)$ 。从 u_i^k 到 u_j^{k+1} 的连接权为 $w_{i,j}^{k,k+1}$ 。设 i_i^k 为 u_i^k 的阈值, i_i^k 的大小可以通过学习进行修改。它的学习规则可以通过增设节点, 转化为和连接权同样的学习方式。

对于输入模式 P , 如果输出层 m 的第 j 个节点的实际输出为 o_j^m , 期望输出为 y_j 的话, 那么平方型误差函数可定义为:

$$E_P = \frac{1}{2} \sum_j (y_j - o_j^m)^2 \quad (3.3)$$

为了减小误差 E_P , 根据梯度下降法, $w_{i,j}^{k-1,k}$ 的修正量 $\Delta w_{i,j}^{k-1,k}$ 为

$$\Delta w_{i,j}^{k-1,k} = - \frac{E_P}{w_{i,j}^{k-1,k}} \quad (3.4)$$

式中, η 为修正系数(>0), 也称学习率、学习系数。

把上式中的偏导数展开, 有

$$\frac{E_P}{w_{i,j}^{k-1,k}} = \frac{E_P}{i_j^k} \cdot \frac{i_j^k}{w_{i,j}^{k-1,k}} = \frac{E_P}{i_j^k} \cdot o_i^{k-1} \quad (3.5)$$

当 $k = m$ 时, 由 E_P 的定义有:

$$\frac{E_P}{i_j^m} = \frac{E_P}{o_j^m} \cdot \frac{o_j^m}{i_j^m} = - (y_j - o_j^m) \cdot f(i_j^m)$$

当 $k < m$ 时, 有

$$\begin{aligned} \frac{E_P}{i_j^k} &= \frac{E_P}{i_j^{k+1}} \cdot \frac{i_j^{k+1}}{o_j^k} \cdot \frac{o_j^k}{i_j^k} \\ &= \frac{E_P}{i_j^{k+1}} \cdot w_{j,i}^{k,k+1} \cdot f(i_j^k) \\ &= f(i_j^k) \cdot \frac{E_P}{i_j^{k+1}} \cdot w_{j,i}^{k,k+1} \end{aligned} \quad (3.6)$$

把上面各式加以整理, 得到连接权 $w_{i,j}^{k-1,k}$ 的修正公式为

$$\Delta w_{i,j}^{k-1,k} = - \eta \cdot d_j^k \cdot \alpha_i^{k-1} \quad (3.7)$$

式中,

$$d_j^k = \begin{cases} (o_j^m - y_j) \cdot f(i_j^m), & \text{当 } k = m \text{ 时} \\ f(i_j^k) \cdot \sum_i w_{j,i}^{k,k+1} \cdot d_i^{k+1}, & \text{当 } k < m \text{ 时} \end{cases} \quad (3.8)$$

$k = m-1, m-2, \dots, 2$ 。

上式中的 d_j^k 可以看作是一般化的误差信号, 它的计算是从 $k = m$ 到 $k = 2$ 递推进行

的。最后一式中的 $w_{j,i}^{k,k+1} \cdot d_i^{k+1}$ 就好像把输出层产生的误差信号 d_j^m 从输出层开始,以相反的方向,加权求和后一层层地传向输入层。这就是这种学习算法之所以叫误差反向传播算法的由来。

如果网络所有节点的作用函数都是 Sigmoid 型的话,即

$$f(x) = \frac{1}{1 + e^{-x}}$$

由于

$$f'(x) = f(x) \cdot (1 - f(x))$$

(3.9a)

因此

$$f'(t_j^k) = o_j^k \cdot (1 - o_j^k)$$

(3.9b)

这是常常使用的一个公式。

在实际应用前面的学习算法时,为了加快收敛的速度,减小振动,可以采用如下的修正公式:

$$w_{i,j}^{k-1,k} \cdot (t+1) = - \eta \cdot d_j^k \cdot o_j^{k-1} + w_{i,j}^{k-1,k}(t) \tag{3.10}$$

其中, η 为指数衰减系数(正数); t 表示学习的次数。

BP 模型虽然从各个方面都有其重要的意义,但也存在着如下问题:

- 从数学上看它是一个非线性优化问题,这就不可避免地存在有局部极小问题,这也是运用梯度下降法所常常遇到的问题。
- 学习算法的收敛速度很慢,通常要几百步、上千步迭代。
- 网络的隐节点个数选取尚无理论上的指导,而是根据经验和多次试验来选取的。
- 对新加入的样本要影响到已经学完的样本,刻画每个输入样本特征的数目也要求必须相同。

3.2.3 Hopfield 模型

1984 年, Hopfield 提出了一种新型的网络模型——Hopfield 模型,如图 3.5 所示:

早在 1982 年 Hopfield 就提出了一种离散的随机模型,而离散的随机模型与连续时间的 Hopfield 模型在许多特性上有很大的关联性。本节只讨论连续时间的 Hopfield 模型,这主要是因为它在硬件上易于实现。

在 Hopfield 网络中,用一组非线性动态公式来描述网络中每一个运算放大器(即神经元)的输出,对第 i 个运算放大器输入的公式化描述为

$$C_i \frac{du_i}{dt} = \sum_j T_{ij} u_j - \frac{u_i}{R_i} + I_i$$

$$v_i = g_i(u_i)$$

(3.11)

其中, C_i 为运算放大器总的输入电容; T_{ij} 为第 j 个运算放大器的输出到第 i 个运算放大器的输入之间的连续权值, v_j 表示第 j 个运算放大器的输出, R_i 为 $g_i(\cdot)$ 确定的电阻值, $g_i(\cdot)$ 为第 i 个运算放大器的 S 型传递函数(假定运算放大器的响应时间可忽略不计), I_i 代表第 i 个运算放大器的外部输入。为使 T_{ij} 具正值或负值,采用翻转或非翻转的运算放大器, T_{ij} 的值为 $1/R_{ij}$, R_{ij} 为连到第 i 个运算放大器的输入和第 j 个运算放大器的输出上的电阻。 R_i 的值由下式确定:

$$\frac{1}{R_i} = \frac{1}{R_i} + \sum_j \frac{1}{R_{ij}} \quad (3.12)$$

式中, R_i 是运算放大器的输入电阻。

图 3.5 Hopfield 连续时间模型

为了弄清 Hopfield 模型的运行, 现考虑把第 i 个运算放大器的输出值看作是 N 维状态矢量的第 i 个分量, 其中 N 为网络中神经元的数目, 状态矢量与式(3.11)一起定性地描述了网络的运行, 网络的状态可看作是在式(3.11)方向上在状态空间里的移动。只要选择内部连接强度的值即可确定状态空间流及其运行的网络特性。针对式(3.11)中描述的系统, 考虑下面的 Lyapunov 能量函数:

$$E = -\frac{1}{2} \sum_i \sum_j T_{ij} v_j v_i + \sum_i \int_0^{v_i} g_i^{-1}(v) dv + \sum_i I_i v_i \quad (3.13)$$

假定矩阵 $T = [T_{ij}]$ 是对称的, 那么 E 对时间 t 的导数为

$$\begin{aligned} \frac{dE}{dt} &= - \sum_i \frac{dv_i}{dt} \sum_j T_{ij} v_j - \sum_i \frac{u_i}{C_i} + \sum_i I_i \\ &= - \sum_i \frac{dv_i}{dt} \cdot C_i \frac{du_i}{dt} \end{aligned} \quad (3.14)$$

由于 $\mathbf{v} = \mathbf{g}(\mathbf{u})$

$$\frac{dE}{dt} = - \sum_i C_i g_i^{-1}(v_i) \cdot \left(\frac{dv_i}{dt} \right)^2 \quad (3.15)$$

式中, $g_i^{-1}(\cdot)$ 是对 v_i 的一阶导数, 由于 $g_i(\cdot)$ 是 S 型函数, 则 $g_i^{-1}(\cdot)$ 是单调递增函数, 即式(3.15)是非负的, 所以 $\frac{dE}{dt} \leq 0$, 而且当且仅当 $\frac{dv_i}{dt} = 0$ 时, $\frac{dE}{dt} = 0$ 。对状态空间的 E 函数,

$\frac{dv_i}{dt} = 0$ 的点即是渐近稳定平衡点, 而且在该点表示网络的最终可能输出集。由于 E

是一有界函数,因而上述公式证明了网络最终是向 E 函数的局部极小点转化;同时由式 (3 .11) 所确定的系统在状态空间中将稳定平衡点随意放置可通过 T_{ij} 的学习来实现。假定希望 n 个正交状态 $\mathbf{V}^s = [v_1^s \dots v_n^s](s = 1, 2, \dots, n)$ 为 n 个存储,则可选取:

$$E = - \frac{1}{2} \sum_{s=1}^n (\mathbf{V}^s, \mathbf{V}^s)^2 \tag{3 .16}$$

在上式中,仅当 $\mathbf{V} = \mathbf{V}^s (1 \leq s \leq n)$ 时, E 取得最小值,且

$$E_{\min} = - \frac{1}{2} N^2 \tag{3 .17}$$

注意此处假定 $v_i \in \{-1, 1\}$, 且 $v_i^s \in \{-1, 1\}$, 将式 (3 .16) 写成式 (3 .13) 的形式且令 $I_i = 0$, 可得

$$T_{ij} = \sum_{s=1}^n v_i^s \cdot v_j^s \tag{3 .18}$$

即只要取 $T_{ij} = \sum_{s=1}^n v_i^s \cdot v_j^s, (s = 1, 2, \dots, n)$, 即为网络的 n 维稳定平衡点向量。

对于由式(3 .11)所确定的神经网络系统,可有如下结论:

- 系统的能量函数沿其平衡轨道单调递减;
- 如果平衡点稳定,则它也是渐近稳定的;
- 渐近稳定平衡点为 E 的局部极小点;
- 系统以大规模非线性连续时间并行方式处理信息,其计算的时间即是系统趋向稳定所需的时间。

3 2 4 Boltzmann 模型

1985 年, Hinton 等人借助统计物理学的概念和方法,提出了一个新的神经网络模型。该模型的概率分布采用了统计物理学中的玻耳兹曼分布,因此,称为玻耳兹曼机。

设网络由 n 个神经元组成,每个神经元具有两个状态,并表示成:

$$x_i = 0 \text{ 或 } 1, \quad i = 1, 2, \dots, n$$

假设神经元之间的结合强度是对称的,即权值

$$w_{ij} = w_{ji} \quad i, j = 1, 2, \dots, n$$

在网络中,当神经元的输入总和发生变化时,将引起神经元状态的更新,这种更新在各个单元之间是非同步的,可用概率分布方法来表示。根据玻耳兹曼机的原理,第 i 个神经元在状态更新时,新状态为 1 的概率是:

$$P(x_i = 1) = \frac{1}{1 + e^{-E_i/T}} \tag{3 .19}$$

式中 T 称为网络的“ 温度 ”,取正数。 E_i 表示第 i 个神经元的状态为 1 和状态为 0 时系统的能量差。

$$E_i = \sum_j w_{ij} \cdot x_j - \theta_i \tag{3 .20}$$

概率 P 的变化曲线如图 3 .6 所示。

由图中的曲线可以看出,在一般情况下,当输入增大时,状态为 1 的概率将提高。同

图 3.6 概率变化曲线

时, 曲线的变化还与温度 T 有关, 温度 T 高, 曲线变化比较平缓, 当 T 趋近于 0 时, 概率 P 的曲线接近于阶跃函数。

上述情况类似于热力学系统中粒子集团的变化规律。当这些粒子集团接上某一热源时, 其最终将达到某一热平衡状态, 这时集团全体的状态概率将符合玻耳兹曼分布:

$$\frac{P}{P} = e^{-(E - E)/T} \tag{3.21}$$

式中, P 表示网络具有 状态的概率, E 表示该状态的能量。由上述公式可以看出, 当温度 $T=1$ 时, 网络两个状态的出现概率的对数差等于两个状态的能量之差。

这样, 当网络中输入部分的神经元被固定在某一特定输入状态时, 网络应能找出最适于特定输入的最小能量结构, 并在某一温度下达到热平衡, 以使网络达到粗水平的能量极小。为使网络达到总体能量极小, 开始时需给一个较高的温度, 然后采用模拟退火技术逐渐降低温度, 以找到更好的能量极小值。

模拟退火算法是近年来新出现的一种随机搜索算法, 其基本思想源于物理学中的退火过程。将模拟退火的方法用于玻耳兹曼机的分析过程, 可利用它求全局最优解, 因为两者具有相似之处, 所以都可以看成是一个寻找最低能量状态的过程。

图 3.7 梯度下降法

图 3.8 模拟退火法

模拟退火算法具有通用性强和解的质量高两个优点。对于传统的启发式搜索算法,例如梯度下降法等,算法每次都是向改进解的方向搜索,这种“贪心”的算法往往导致算法只能找到一个局部最优解,而不是全局最优解,如图 3.7 所示。模拟退火算法可克服这个缺点,它在每步搜索中不但可以往改进解的方向走,而且还允许以一定的概率往恶化解的方向走,即它不但能接收好的状态,而且还以玻耳兹曼概率分布接收差的分布。这种做法带来的好处是使算法在一定条件下可以收敛到全局最优解,如图 3.8 所示。

玻尔兹曼机(BM 网络)具有输入层、输出层和隐含层。这一点与 BP 网络相似。但是 BM 网络层内神经元彼此联系,这一点不同于前馈神经网络。玻尔兹曼机(BM)的一般结构可以表示为图 3.9。

BM 神经网络虽然可以在一定条件下获得全局最优解,但是,为了求得网络状态的最优解需要采用非常缓慢而且细致的退火搜索过程,这就需要耗费大量的时间;为了求得最佳的网络权系数又需要非常多的学习周期来进行学习,并且每个周期中需要用很多次运算来进行统计计算,以保证统计结果的正确性。这一切都限制了它的实用性。

图 3.9 玻尔兹曼机的一般结构

3.3 神经网络学习规则

神经网络中的神经元必须有机地结合在一起,按照某种规律运动,才能体现出某些“智能”特征。这就要求神经元之间的连接方式和权系数必须按一定的规律安排。又由于神经网络一般都很大、很复杂,它的连接方式和权系数不可能事先人为地设置好,因此,神经网络必须能自学习或自组织形成所需要特性的结构。

神经网络的学习可以定义为:学习是状态的变化,以达到性能上有意义的稳定性。这种状态变化是由反复作用、训练或实验而得到的。学习规则是指神经网络学习的一套精确算法,用以确定神经元应如何改变它的连接权重,从而正确响应训练。目前已有许多种学习规则,下面我们介绍几种常用的学习规则。

3.3.1 Hebb 规则

40 年代末期,D .Hebb 首先提出了一种神经网络学习算法,称为 Hebb 规则。以此规则为基础,出现了多种形式的学习算法。在介绍 Hebb 规则以前,我们先介绍它的基础——Hebb 定律。

Hebb 定律是相联式学习的基本心理学定律,它既是心理学的基本定律,也是人工神经网络的基本定律。Hebb 定律的全部含义是:一切联想(长期的)记忆都保存在中枢神经系统的突触连接中,而且所有学习均意味着突触耦合强度的变化。Hebb 规则就是在 Hebb 定律基础上提出的。假定第 j 个神经元到第 i 个神经元的连接强度(权系数)为 w_{ij} , Δw_{ij} 表示权系数的变化, a_i^p, a_j^p 表示神经元 i 和 j 在输入样本 $P(p = 1, 2, \dots, p)$ 时的活化状态,则 Hebb 规则可用下式表示:

$$w_{ij} = \frac{1}{p-1} \cdot a_i \cdot a_j \quad (3.22)$$

$$w_{ij} = \frac{1}{u=0} a_i^u \cdot a_j^u \quad (3.23)$$

式中, η 为学习速度系数。

Hebb 学习规则的优越性就是表现在神经元 i 和 j 的激励值 a_i, a_j 的相关性, 即能产生有用的关联学习。对于某一特定的神经元 i , 当其处于活化状态时, 它将倾向于激励那些曾与其相关的活化神经元。

3.3.2 Delta 规则

最为流行的神经网络学习算法是 Delta 规则, 它是 Hebb 规则的延伸, 是由 Widrow 和 Hoff 在 1960 年提出的。Widrow 和 Hoff 用自适应线性元件将输出值和输入值进行类比并用权系数的改变来产生真实的、希望的输出值, 其主要目的是减少均方差信号。为了减少真实输出值与希望输出值的误差, 我们需要调整权系数。调整权系数的规则即 Delta 规则, 可写成下述公式:

$$\Delta w_j = (d(p) - y(p)) \cdot I_j(p) = \eta \cdot (p) \cdot I_j(p) \quad (3.24)$$

Δw_j 表示对第 j 个权值的变化量。其中 $d(p)$ 为第 p 个输入值的期望输出值, $y(p)$ 为第 p 个输入值的估算输出值, $I_j(p)$ 为第 p 个输入值的第 j 个分量, η 为学习率。

为验证 Delta 规则是梯度下降法, 可以证明每个权值误差的导数与权值变化成反比, 假设

$$E(p) = \frac{1}{2} (d(p) - y(p))^2 \quad (3.25)$$

是对第 p 个值误差的度量, 而且把 $E = \sum_p E(p)$ 作为总误差, 那么, 用链式规则对第 j 个权的 $E(p)$ 导数为

$$\frac{\partial E(p)}{\partial w_j} = \frac{\partial E(p)}{\partial y(p)} \cdot \frac{\partial y(p)}{\partial w_j} \quad (3.26)$$

上式中右端第一项是估算输出值的误差变化, 第二项是对第 j 个连接的估算输出的变化量, 从公式(3.25)有

$$\frac{\partial E(p)}{\partial y(p)} = - (d(p) - y(p)) = - \delta(p) \quad (3.27)$$

因为 $y(p) = \sum_j w_j I_j(p)$, 公式(3.26)的第二项是

$$\frac{\partial y(p)}{\partial w_j} = I_j(p) \quad (3.28)$$

将公式(3.27)和(3.28)代入公式(3.26)得

$$-\frac{\partial E(p)}{\partial w_j} = \delta(p) \cdot I_j(p) \quad (3.29)$$

将公式(3.24)与公式(3.29)进行比较, 表明 Delta 规则对每个连接权沿着均方差的负导数方向改变权值。因此, 得

$$\Delta w_j = - \frac{\partial E}{\partial w_j} = \delta(p) \cdot I_j(p) \quad (3.30)$$

所以,Delta 规则实现了在误差度量 E 上的梯度变化。

3 3 3 ART 规则

自适应共振理论(adaptive resonance theory,简称 ART) 是一种非指导式学习规则,靠自组织途径响应输入模式序列,用区分输入信息的特征对输入模式进行分类。

ART 模型中一般有三层,其中输入层和输出层称为短期记忆层,中间层称为长期记忆层,输出层到输入层存在一个反馈矢量,称为由顶向下的自稳机制。长期记忆层记住已学到的东西,已学模式以权重方式存储在长期记忆层中。短期记忆层则把握当前的输入模式、模式的分类和所期望的模式。自顶向下的自稳机制是为了保证长期记忆层中的记忆具有足够的牢固性。每当一个模式输入到系统,所期望的模式便从长期记忆中调出比较,如果输入的模式与所期望的各模式之一相类似,网络就会把输入模式分到对应类中;如果没有相似的,则新的一个输入矢量表示的类别便形成了,并作为新一类首次记忆保存。

ART 的非指导学习方法又叫“ 竞争学习 ”。这项技术是让神经元为一个有限量的活性值而相互竞争,最简单的竞争学习是所谓“ 胜者取全部 ”的规则,即最大输入的神经元被激活,其余的神经元被抑制。

自适应共振理论使用分类法来选择输入模式的类。输入模式激活短期记忆,并通过长期记忆来处理,具有最高活性的神经元被选中为“ 优胜者 ”。当给出一个输入模式选中一个优胜者时,网络就通过反复调整长期记忆的权重的办法来处理输入模式,并将输入模式的特征加到长期记忆之中。按照上面给出的规则,只有当新的输入模式与存入记忆中的某个老的模式足够相似时,两者才能融洽,即有关的权重系数进行调整,从而使长期记忆得以改变。这就造成了一种自适应谐振状态,这就是 ART 名称的由来。

ART 规则是一个比较接近人脑处理信息方式的学习规则,它可以“ 实时学习 ”,即学习和工作是分不开的。这种学习能够达到稳定、可靠的结果,这种学习是自治和自组织的,学习过程无需教师指导,因此是一种无监督学习。这种学习可以形成“ 集中注意 ”的状态,也就是说可以有选择地学习,因而具有较高的学习效率。这种学习还可以完全避免陷入局部最小点问题。

3 3 4 Kohonen 规则

Kohonen 规则也采用竞争学习的概念,是一种无监督的学习规则。

设神经元 i 从别的神经元获得了许多输入 x_{ij} , 这些输入构成了一个矢量 $\mathbf{x} = \{ x_{i1} \quad x_{i2} \quad \dots \quad x_{in} \}$ 相应于每一个输入 x_{ij} 的权重为 w_{ij} , 这些权重也可以视为一个矢量 $\mathbf{w} = \{ w_{i1} \quad w_{i2} \quad \dots \quad w_{in} \}$ 。当神经元的连接函数是求和函数时,其结果为输入矢量与权重矢量的点积: $\mathbf{x} \cdot \mathbf{w}$, 它是这两个矢量之间距离的量度。

Kohonen 规则也使用了“ 胜者取全部 ”的策略,而胜者就是权重矢量 \mathbf{w} 与输入矢量 \mathbf{x} 最近的神经元,或者说是活性最大的神经元。算法将调节优胜者神经元的相应权重,形成一个新的矢量,使之更靠近输入矢量。这样, Kohonen 学习规则对于外界刺激最强的区域是

$$\mathbf{w}(t+1) = \mathbf{w}(t) + (\mathbf{x}(t) - \mathbf{w}(t))$$

其中 η 为学习常数。自动调节,直到权重向量与输入向量的某一最大主分量重合为止。

可见,Kohonen 学习规则是一种无指导教师的聚类方法,它将所形成的聚类中心映射到输出平面上,而保持拓扑结构不变。这是一种两层网络,在输出层神经元上的输入量是输入矢量 \mathbf{x} 与权重矢量 \mathbf{w} 的点积。

3.3.5 Boltzmann 规则

Boltzmann 规则是一种加强式学习规则,用一个目标函数来估价网络的输出。与其它学习规则相比,Boltzmann 学习规则独特之处在于它用了概率函数来改变连接权重。

我们知道,所谓学习就是改变神经网络中各个权重系数的数值。系统的能量函数为

$$E = - \frac{1}{2} \sum_{i,j} w_{ij} \cdot x_i \cdot x_j$$

能量 E 所具有的状态就是由这些权重系数 w_{ij} 决定的,因而改变权重系数就会导致 E 的变化。由式(3.21)可知, E 的变化直接引起系统中各种状态(尤其是其中的各个能量极小点)出现概率的变化。我们知道,系统可以在三种情况下运行:第一种情况,系统与外界接触的神经元全部受到约束,这种全约束情况相当于系统的“学习情况”;第二种情况,与外界接触的神经元中有部分受到约束,这种部分约束情况,与系统的“工作情况”相对应;第三种情况,所有神经元都不受约束,这种情况称为“非学习情况”或称为“自由状态”。学习准则是,在学习情况下调整各个系数,使得对于各种特定的约束而言,在学习和工作两种情况下,系统通过搜索能够达到一致或非常接近的热平衡状态。事实上,这些状态就是系统能量函数中的一些极小点,通过学习后,它们都具有某种特定的出现概率。在自由状态下,系统通过搜索达到上述状态也应具有一定的概率。

在 Boltzmann 学习规则中,神经网络中神经元状态被视为类似某种物质体的微观状态,而温度 T 是改变网络权重的参数,温度高,则权重改变的概率大,温度低则引起改变的概率小。这种学习规则一般包括以下几个步骤:

将所有的网络权重随机化,并设初始温度为高温 T ;

用输入矢量和网络输出值计算目标函数的值,即计算输出值是否符合目标;

根据 Boltzmann 分布 $P(x) = e^{-x^2/T^2}$ 随机地改变每一权重;

重新计算目标函数;

如果目标函数已得到改善,就将上述权重改变确定下来。如果目标函数未得到改善,则一般不确定,降低温度 $T(n+1) = T(0) * (1/\log(1+n))$ 回到步骤 。

温度在网络初始化时高,权重改变概率大;随着网络被训练而降低温度,使权重每一次改变量越来越小,这样,网络可趋向最佳状态。目标函数起到引导网络向最佳状态逼近,正确的方向得到加强,故有人称之为“加强式”学习。又因为权重改变式是一个概率函数,故有人称这种网络是“随机网络”。

3.4 基于神经网络的机器学习

3.4.1 机器学习

人工智能领域中的许多尚待解决的问题都与机器学习直接相关,机器学习的实现将使人工智能理论及其应用的研究产生较大的进展。目前,机器学习已成为人工智能的瓶颈问题之一。

1. 学习的意义

学习的含义变化无穷,给学习这个概念下定义如同给知识下定义一样困难。神经学家把“学习”看成是“条件反射的形成过程”;心理学家把“学习”定义为“学习是系统积累经验以改善其性能的过程”;“学习是个体经验的获得以及行为变化的过程”;控制论学者认为“一个系统,若在它的环境发生变化后的时间 T 内,相对性能函数 P 的响应是令人满意的,就可以认为是学习系统”;人工智能研究者提出“任何系统,只要它能够从某个过程或环境的未知特征中学习及有关信息,并且能以学到的信息作为经验用于未来的估计、分类、决策和控制,以便改进系统的性能,那么这就是学习系统”;M. Minsky 对学习下了一个简单又易于理解的定义:学习就是使我们的大脑进行有用的变化。总之,学习必然要与环境相互作用,不断积累系统所拥有的知识,目的是使系统的工作能具有更高的智能水平。

学习又可根据其特性区分为学习者技能的提高和知识的获取。技能提高在生理方面可解释为建立和加强一个神经网络模式,使其完成所需的功能;知识获取则涉及到对知识进行收集、分类、组织、集成和抽象。

2. 学习策略的分类

对学习策略进行分类的一个重要尺度就是要求学习者进行了多少推理。推理或归纳推理都涉及到生成规则或将一些特殊的实例一般化。一个广为接受的学习策略分类法就是从不需要推理到需要高度推理来进行分类。

(1) 机械式学习 无论是人还是机器,这都是最低级的学习。在人类学习中,对应记忆一个事实,而该事实不会导致推理发生。在机器学习中,对应通过编程或将知识直接存放在数据库系统中,在使用它们时不需要做进一步的处理或翻译。在使用机械学习方法时,如果能“忘记”那些很少使用的知识,或已证明是错误的知识,将会提高系统的效率。

(2) 通过指导学习 学习系统将从教师或书本上获得的知识转换成可直接使用的内部表示的知识。这种学习模式与课堂上教师和学生的关系非常相似。在此学习模型中,学习者所承担的工作是对知识源所提供的知识进行语法重构。

(3) 通过演绎学习 这一级的学习比通过指导学习更加强调知识的转换,对知识源中知识的表达形式的限制放宽了。学习者使用这些知识进行演绎推理,并将它们转变为包含了所有原始信息的另一种有用的结论。演绎学习包括重构、编译和组织三个过程。

(4) 通过类比学习 该学习模型结合了通过演绎和通过归纳两种学习方法。首先,进行归纳推理,在问题领域和存储在学习者知识库中的相似领域之间寻找共同的子结构;其次,将所选择的相关领域的结构映射到问题领域中。在映射之前需要对此共同子结构进行一般化,再使用演绎逻辑来完成这一映射。

(5) 通过归纳学习 该学习模型要求大量的归纳管理。通过归纳学习的方法又可分为通过实例学习、通过观察学习和通过发现学习。

通过实例学习涉及到概念获取的过程,该过程通过对教师环境以及学习者自身的知识库提供的一组实例进行推理,获取一般概念的描述。所获取的概念一般能解释所有的正实例,严格地不包括任何负实例。如果实例的提供者是环境,系统的工作应像实验一样允许退步。

通过观察和通过发现学习不需要教师的帮助,它们都是通过推理获取一般规则和规律,用它们来解释所观察到的问题。这些方法所使用的推理技术包括数据填充、对象的分类和概念分群等等。

3 . 一个机器学习的模型

一个机器学习的模型如图 3 .10 所示,该模型由以下四个部分组成:

(1) 知识库 它是规则或数据结构的集合。知识库是一个动态的结构,它将随着知识的增减和修改而变化。正确选择知识表示方法是机器学习成功实现的关键。

(2) 一个任务模块或行为算法 它使用知识库中存储的知识和用户临时提供的知识对当前的问题进行求解。任务模块必须产生一个量化的输出,用于测试系统行为与理想系统之间的误差。

图 3 .10 一个机器学习的模型

(3) 一个后退模块或循环 它将任务模块的输出与一个理想系统进行比较。后退模块提供的信息是一种报错信息,它指导学习者修改知识库。

(4) 学习模块 它使用报错信息对知识库中的规则进行修改,减少实际输出与理想系统正确输出之间的差距。

这个系统表面上看起来似乎很简单,但要真正实现一个实用的学习系统要考虑到许多因素。其中最重要的因素就是知识表示,要易于在机器中实现,要与真实世界中的知识有直观对应关系,要灵活易变,要支持知识库的动态修改等等。

3 4 2 基于神经网络的机器学习

学习问题可以看成是获取一个将输入矢量(激励)映射到输出(响应)的函数的过程。为了学习这个映射函数,学习系统接收一组训练事件,由此构造映射函数的内部表示。从这一角度看,神经网络的学习与传统系统中的学习没有区别,然而,在如何完成学习方面,却有很大的差别。主要的区别涉及到知识的表示和知识的获取两个方面。

1 . 神经网络的知识表示

(1) 神经网络知识的结构 在神经网络模型中,我们使用概念、属性及属性值来描述客观世界。属性值(如 red)、概念值(如 apple)和单个神经网络块统称为一个知识单元。属性和结构链(如 is-part-of)称为特性。概念是特性和值的集合。特性的值也是概念。因此,概念可以是任意复杂的。

一般情况下,概念可以分为事例和类型两类。通常人们可能把外部世界看成是由事例组成的,一个事例的特性和值集合对应概念结构中的事例,而类型是对事例的抽象和概括,它使人们可以把握复杂的客观世界。特性也可以分为两大范畴:属性和结构链。结构链在神经网络知识表示和推理之间架设桥梁,这恰好反映了认识论的观点:世界知识是高度组织的。利用这些组织结构可以形成有关推理的有效策略,每一结构链体现出一种组织策略。如果我们将知识单元看成一个数据结构,那么这些结构链便是数据结构之间的连接弧,这些弧为推理提供了有效的路径。

在神经网络知识表示语言中,结构链定义为知识单元之间变量(属性)的约束关系。在神经网络推理过程中,知识单元之间的变量约束关系导致了神经网络知识库中类型之间的等级结构。一般来说,我们可以在相同的事例上定义多个等级结构,这是因为一个事例可以和神经网络知识库中的多个知识单元进行变量约束,多个等级结构的结果是一个概念在层次结构上可以有多个祖先。

(2) 神经网络知识表示的结构 一个神经网络块就是对一组事实进行归纳,它相当于传统智能系统的一组规则,所以,神经网络知识库也可以看成是由特殊的事实和规则组成。但是它与普通的知识库有本质的不同,在这个知识库中,不需要一个独立的解释器,网络中的联结、连接线上的权值以及网络本身的计算特性不仅对知识本身编码,而且对神经网络知识库中各部分知识在计算过程中的相互作用进行编码。从语义的角度来讲,这样的规则是隐式的。

2 . 神经网络的知识获取

知识获取是人工智能或知识工程研究的瓶颈问题之一。这个问题一方面体现在人类的知识很难用一些规则来描述,甚至很难用简单的语言来描述;另一方面则是由于不良结构所产生的知识组合爆炸问题。而在基于神经网络的并行分布式模型中,知识获取是在给定输入模式和输出模式下,通过学习过程自动地调节各层之间的结合权值和隐含节点的阈值来完成的。与传统的知识获取不同,在这样一个网络中,理论上可拥有极大量的知识,如果神经元有 N 个输入单元,输入模式是二值逻辑,则可提供表达知识的样本数为

2^N 。神经网络在一定程度上模拟了专家凭直觉来解决局部情况的不确定性问题。当输入模式接近于学习过程中的训练样本输入时,网络的推理结果也表示出接近样本输出的倾向。更主要的是,神经元以连续的方式进行学习,当输入与输出不匹配时,可以在原有基础上继续学习,提高学习精度,并在原有基础上增加新的训练样本,继续学习产生对输入而言合乎要求的输出。

用神经网络来进行知识获取,需要解决的首要问题是给网络结点赋予一定的物理意义。因此,我们需要建立背景库,将特定的属性、特征或概念与神经网络的输入、输出节点对应起来,把一个背景库下的事例放入同一事例库中进行学习。对于分类型智能系统,主要通过有导师学习算法,通过典型事例进行学习,而对一些类别不清楚的事例,则采用无导师学习算法。

由于 BP 算法具有很强的映射能力,它已在模式识别、知识获取、联想存储等方面获得了一定的应用,并已发展得比较成熟,因此,我们选择 BP 网对典型事例进行学习。

BP 网络是一种层次型神经网络,它由输入层、输出层和若干隐含层组成。其学习原理符合行为学的特点,即人在实践中一边学习,一边修正错误,再学习,再提高,直到其行为达到预期的目的。因此,BP 算法把学习分为两个阶段:

(1) 输入值正向传播阶段 输入值从输入层传播节点,经作用函数处理后,输出信息传播到下一层隐含节点,最后达到输出层,经作用函数处理后成为网络的输出,这就是网络的正向传播过程。

当给定一个输入样本后,在第 k 层第 i 个节点得到的输入 U_i^k 可用下列公式表示:

$$U_i^k = w_{i,j}^{k,k-1} \cdot I_j^{k-1} - \theta_{ki}$$

这里, $w_{i,j}^{k,k-1}$ 是第 k 层第 i 个节点与第 $k-1$ 层第 j 个节点之间的连接权值, θ_{ki} 是第 k 层第 i 个节点的阈值, I_j^{k-1} 是第 $k-1$ 层第 j 个节点的输入。在输入层 ($k=1$), I_j^k 就等于网络的输入;当 $k>1$ 时, $I_j^k = f(U_j^k)$ 。 f 是 S 型函数,通常 $f(x) = \frac{1}{1 + e^{-x}}$ 。

(2) 误差反向传播阶段 用函数 $E_p = \frac{1}{2} \sum_j (y_j - o_j^m)^2$ 来度量误差。在第一个阶段结束后,可以得到一个误差值,若该值不满足事先给定的精度,则表示对该批样本的训练还未获得理想的效果。为了使用网络的输出符合目标输出,必须从输出层开始,逐层调整权值和阈值。BP 网络的误差修正公式为:

$$w_{i,j}^{k-1,k} = - \delta_j^k \cdot \alpha^{k-1}$$

可以证明 BP 网络使用梯度下降结束修正误差。

BP 网络具有很强的映射能力。有人已经证明,对于任意给定的 $\epsilon > 0$,则对于任意一个二阶连续函数 f ,存在至少一个三层 BP 网络,它使得 f 的平方差小于 ϵ 。

虽然如此,由于 BP 网络中隐含节点的选择无规律可循,因而我们未必能构造一个能实现所期望映射的恰当网络;其次,由于梯度下降法收敛速度慢,影响 BP 网络走向实用,再者,梯度下降法只保证函数收敛到局部极小点,而不保证是全局极小,一旦网络陷入局部极小,整个学习过程就将失败。这三个问题不解决,就难以成功地用 BP 算法进行知识获取。为此,人们提出了许多改进 BP 算法,同时,也给出了可用于知识获取的其它一些

神经网络模型,有兴趣的读者可参阅其它有关文献。

3.5 神经网络专家系统

3.5.1 一般描述

虽然专家系统把它的问题领域限定于某一特定的专门领域,避免了通用型人工智能系统所固有的矛盾,即问题的无限复杂性与系统处理和理解能力的有限性的矛盾。专家系统已经在不少专门领域表现出相当出色的工作能力,在许多场合不仅达到甚至有的还超过了人类专家的工作能力,从而使基于“功能主义”软件方法的人工智能理论走出了积木游戏世界,逐步进入现实世界。虽然专家系统的技术仍在不断发展不断完善的阶段,新的不同领域的专家系统如雨后春笋般被建造出来,但是,专家系统技术本身固有的问题和局限性并没有因此而完全解决,随着专家系统技术更深入的研究,这些问题和局限性也日益明显地暴露出来。

在开发专家系统的过程中,专家系统所面临的问题主要有:

1. 知识获取问题

专家系统本质上是一个知识工程系统。知识获取、知识表示、知识存储和知识运用构成了专家系统工作的主要内容,其中知识获取是首要的前提。但在建造专家系统时,普遍存在着知识获取比较困难这一难题。专家系统的知识来源于领域专家,为了将领域专家的知识总结归纳出来,只有两条途径:即由知识工程师向领域专家获取或让机器自动获取。如果采用第一种方法,由知识工程师向领域专家“求教”,那么,在很多情况下,我们经常遇到的问题是领域专家自己也很难把自己的知识讲清楚;一些专家在很多场合下可能知道怎样去解决他所要对付的问题,但是说不出为什么应当这样去解决问题,因为他们有时只是凭经验或直觉去这样做的,何况有时专家解决问题的“知识”未必很好,甚至也可能有误。这就给知识获取本身带来难题。如果知识是由机器自动向“专家”学习得到的,由于机器学习能力目前仍相当低下,这样,专家系统可自动获取知识的能力也就受到很大很制。

2. 知识的“高台效应”问题

专家系统所面临的另一个难题是其性能的“高台效应”问题。知识的运用都有一定的适用范围,由于专家系统的专业领域相对比较狭窄,其“良好”定义的问题及其处理的问题知识体系也有一定限度,所以对完全“匹配”的问题,专家系统能以专家水平做出处理并给出可靠的解答,而对不完全“匹配”的问题,或超出专家系统知识领域的问题,系统可能非但不能正确处理,有时还会“神经错乱”,给出“不伦不类”的答案。麻烦的是,专家系统本身往往并不能判断什么时候或什么情况下问题已接近或超出了它的边界。

3. 知识存储容量与运行速度的矛盾问题

专家系统所面临的还有一个难题就是知识的存储容量与运行速度的矛盾问题。为了

能更好地处理领域问题,我们希望知识越多越好,但这又给系统的运行带来了一些困难。由于现有的 Von Neumann 计算机只能支持串行的处理方式、集中处理的方式以及严格化的符号处理方式,中央处理器与存储器之间狭窄的数据通道与大规模的搜索和匹配要求发生了矛盾,以至不得不面对这种现象:在专家系统中,知识越多,解决问题越慢,这从根本上限制了专家系统解决大型复杂问题的能力。

所有这些困难和问题促使人们冷静思考,人们越来越意识到,单纯依靠基于 Von Neumann 计算机从功能定义和软件方法来解决人类智能的复制与模拟的问题是不现实的、不科学的。为了进一步提高专家系统的性能,有必要“引入”基于“联接机制方法”的神经网络技术。

3 5 2 基于神经网络的专家系统

神经网络系统最主要的特征是大规模并行处理,信息的分布式存储,连续时间非线性动力学,全局群体作用,高度的容错性和自组织、自学习及实时处理。它可以直接输入范例。信息处理分布于大量神经元的互连之中,并且具有冗余性,许许多多神经元的“微活动”构成了神经网络总体的“宏效应”。这些正是与传统人工智能的差别所在。人们可以利用神经网络系统的自学习功能、联想记忆功能、分布式并行信息处理功能来解决专家系统中的知识表示、知识获取和并行推理等问题。

分布性是神经网络之所以能够解决专家系统中知识获取这个瓶颈问题的关键所在。与传统计算机局域式信息处理方式不同,神经网络是用大量神经元的互连及对各连接权值的分布来表示特定的概念和知识的。在进行知识获取时,它只要求专家提供范例(或实例)及相应的解,通过特定的学习算法对样本进行学习,经过网络内部自适应算法不断修改权值分布以达到要求,把专家求解实际问题的启发式知识和经验分布到网络的互连及权值上。对于特定的输入模式,神经网络通过前向计算,产生一个输出模式;其中各个输出节点代表的逻辑概念同时被计算出来。特定解是通过输出节点和本身信号的比较而得到的。在这个过程中,其余的解同时被排除。这就是神经网络并行推理的基本原理。在神经网络中,允许输入失真的学习样本。只要输入接近某一学习样本的输入模式,则输出亦会接近学习样本的输出模式。这种特性使得专家系统具有联想记忆的能力,这就是神经网络专家系统的基本原理。

神经网络专家系统的目标是利用神经网络的学习功能,大规模并行分布处理功能,连续时间非线性动力学和全局群体作用实现知识获取自动化;克服“组合爆炸”和“推理复杂性”及“无穷递归”等困难,实现并行联想和自适应推理;提高专家系统的智能水平及实时处理能力。神经网络专家系统的基本结构如图 3.11 所示,其中各模块的功能是:自动知识获取模块研究如何获取专家知识;推理机制提出使用知识去解决问题的方法;解释模块用于说明专家系统是根据什么推理思路作出决策的;I/O 系统是用户界面,它提出问题并获得结果。

知识获取包括提出所需神经网络的结构(包含网络层数、输入、输出及隐含节点个数);组织待训练的学习样本;使用神经网络学习算法,通过对样本学习,得到所需权值分布,从而完成知识获取。

图 3.11 神经网络专家系统的基本结构

知识库由自动知识获取得到,它是推理机制完成推理和问题求解的基础。知识库可以不断创新,即在原基础上对新样本进行学习,获得具有更多知识与经验的新的网络参数。

神经网络专家系统的推理机制与传统的专家系统所用的基于逻辑的演绎方法不同,它的推理机制为一个数值计算过程,主要由以下三个部分组成:

(1) 输入逻辑概念到输入模式的变换,并根据领域的特点,确定变换规则,再根据相应规则,将目前的状态变换成神经网络的输入模式。

(2) 网络内的前向计算:根据神经元特征,其输入为 $x_i = \sum w_{ij} \cdot y_j$, w_{ij} 为连接权值, y_j 为神经元输出,且有 $y_i = f_i(x_i + \theta_i)$,其中 θ_i 为神经元的阈值, f_i 为单调非线性函数。通过上述计算即可产生神经网络的输出模式。

(3) 输出模式解释:随着领域的不同,输出模式的解释规则各异。解释的主要目的是将输出数值向量转换成高层逻辑概念。

下面,我们看一个“选择工业装载车的专家系统”的例子。

选择合适的装载车涉及到:

弄清可供选择的装载车的特点;

确定应用范围;

用合适的装载车适应所关心的应用范围,以便使选择的装载车足以完成应用范围中所需的活动。

为了弄清装载车的特点,需要对可用的装载车进行分类。分类必须根据装载车的特征,因而工业装载车可以按下列因素分类:

装载类型,例如前装或侧装;

轮胎类型,例如半实心轮胎或充气轮胎;

发动机类型,例如内燃机或电动机;

装载车的载重量。

按照这种分类,商品化装载车可以分为下列几类:

载重量可达 3 吨,半实心轮胎,内燃机动力;

载重量可达 4.5 吨,前装,气动平衡装载车;

载重量 8 吨,侧装,平衡跨架式装载车。

这三种类型的商品化装载车就可以作为神经网络训练之用。

选择合适的装载车可用的特性和要完成的作业。例如,侧装的平衡装载车对于整装仓库的装卸是一件合适的选择,因为这种装载车可以在很窄的通道中拐弯,并且能将普通尺寸的整装货物举升至一定的高度。由于整装仓库中常用窄通道,因而这种选择看来是合理的。

基于神经网络的选择工业装载车的专家系统,采用一个具有一层隐含层的 BP 网络,其训练是由专家安排的,其输入模式为:

$\mathbf{x} = \{$ 货物仓库的通道宽度和长度;
货物装卸的作业高低;
对于前装或侧装的要求;
单件货物的最大尺寸;
仓库地面状况;
装卸货物的效率。 $\}$

输出模式为

$\mathbf{y} = \{$ 装载类型;轮胎类型;发动机类型;装载车的载重量;装载车的尺寸。 $\}$

按照一定的规则,可以将上述模式全部数值化。训练好的神经网络实现从 \mathbf{x} 到 \mathbf{y} 模式的映射,也可以说,神经网络学会了从输入条件到输出结论的某种推理。

对于神经网络专家系统,我们不需要建立如何从输入模式 \mathbf{x} 到输出结论 \mathbf{y} 之间的推理规则。神经网络从例子(或训练集)中学习并建立起它特有的推理模型,从而完成这一知识的表达。这一点十分重要,即神经网络从特定事例推广到更一般性原理(规则),它不仅记住了已给实例,而且抓住了其特征。这些特征可以用来区分已知实例,还可以用来推理、分析未知实例。

第 4 章 智能制造环境下的 产品模型及其建模

智能制造的目的是运用集成知识工程技术、制造软件系统技术、机器人视觉和机器人控制技术,归纳技工的制造技能与专家知识,对生产过程进行建模,使智能机器能够在没有人实时干预的情况下进行自动化生产。

现代工业生产作为一个有机整体不仅表现在各制造环节之间存在的技术型联系,而且还表现在人类专家的制造智能的统一体特性方面。制造智能理论及处理技术就是要研究整个制造环境中的种种智能源的开发、描述、集成、共享与处理,特别是各种制造环节的知识获取(knowledge acquisition, KA),而制造过程的描述与建模是实现智能制造的一个关键环节。本章将介绍产品模型及与其有关的模型和建模方法。

4.1 产品模型

随着制造过程各方面自动化水平的提高,设计和生产需要有一个统一的产品模型,以提供制造过程各方面所需的信息,在智能制造环境下,要求计算机对产品的建模能够表示产品的全部信息。它类似于传统制造过程中的产品图纸,只不过是由计算机模型化了。然而,现代制造技术的研究和开发,其本质是要求进一步提高产品信息的智能化处理水平,从而研究智能制造环境下的产品模型已经成为必然趋势。

4.1.1 产品模型的概念

模型是用来表示实际的或抽象的物体和现象,它给出被处理对象的结构和性能,并产生其图形。建立模型是对被处理对象进行计算、分析、模拟、研究的一个基础。

模型本身并不一定要计算机化,但这里我们要讨论的是计算机模型,即在计算机上建立一个对象的结构和性能的表现形式,并通过对这种表示形式的操作来解决有关问题。在传统制造过程中,表达媒介是工程图纸。在以计算机为基础的现代制造系统的制造环境下,表达产品数学模型的媒介就要求由计算机(其存储媒介)来代替。这种代替不是一种简单的替代,而是从一个方面蕴含着制造系统发展的一场深刻革命,也就是朝着高度自动化的智能制造系统发展。

对自动化智能制造系统而言,在控制理论和方法中,建模问题可以分为两大类:对被加工零件的产品建模问题以及加工过程中的过程模型问题。加工过程模型有刀具磨损模型、切削成型模型、加工质量数学模型等等。传统的建模方法有:根据物理原理推导出模型结构和根据实验和经验选取各种参数或修正系数,这对于复杂的制造过程建模来说并不是很有效的。现代采用时域或频域分析建模方法。由于制造过程复杂,过程模型很难

用解析的方法表达为准确的数学模型形式,所以往往采用智能建模即神经网络的建模思想。现有的模糊系统、灰色系统、学习系统的方法,也可对复杂非线性的、时变的系统建立其动态行为模型。

本章对制造过程建模不作介绍,而主要针对被加工零件的产品模型进行讨论。

产品模型是对具有某种功能的产品(已生产的实际产品或将要生产的产品)在三维欧氏空间,建立起它的数学模型,然后在某种媒介上表示这种数学模型。要求在这个数学模型里描述的任意一个产品,满足这个产品的功能要求,并且能够为制造过程所生产;同时也尽量要求其满足功能和生产要求的一切产品为这个数学模型所描述。这样,这个数学模型所表达的文件就成为制造过程的许多任务所依据的产品描述文件,从而要求建立一个既能反映产品生命期中各阶段的数据要求,又能反映各阶段数据关系的统一的产品模型。

产品模型是产品智能制造应用的核心,它包括了产品的定义信息,与产品设计和制造有关的技术、管理信息等,其中既有形状信息,又有非形状信息。

分析产品零件的信息体系结构是建立一个共享的产品模型的首要任务。为适应集成知识工程、并行工程以及智能制造的要求,必须做到:

(1) 产品零件信息模型应是一个动态的数据结构,便于设计者动态操作。设计者可以在设计过程中的任何阶段将设计中间结果(不完备的零件信息模型)移交工艺评价,并根据评价结果,修改模型或继续设计。

(2) 产品零件信息模型应强调零件功能与零件特征的关系以及零件功能与各特征功能的关系,并在各种关系之间建立一种映射联系。零件设计与工艺设计在满足零件功能要求的前提下进行交叉设计,使设计出的产品既满足指定的功能,又具有经济的可制造性。

(3) 产品零件信息模型应具有数据、几何和知识,便于多知识源的协同处理,可采用黑板结构,在系统“管理者”的协调控制下,对领域黑板上的当前零件信息模型进行操作。

4 1 2 产品模型的演变

1992 年日本京都大学冲野教郎教授在我国讲学时,提到产品模型的发展过程,是从二维模型、三维线框模型、曲面模型、实体模型、特征模型、产品模型到最新的生物模型。可见实体模型是在成熟的二维模型、三维线框模型、曲面模型基础上进行研究和发展的,它是实现三维几何实体的完整信息表示的理论、技术和系统整体的概括。通过这种完整的和无二义性的三维几何实体的表示,可以自动地算出被表示物体的任何有效定义的几何特性。冲野先生还认为,造型优美的线框边界往往是产品外形设计的前提,作为产品模型,在零件的三维几何描述中往往需要引用中心线、基准面、局部坐标系等,这就需要 CAD 系统可以兼容线、面、体于一个模型中,即产品模型。

近几年来,对产品信息的描述、产品建模方法的研究,为智能制造技术的发展奠定了基础。

1. 二维模型与工程图绘制

二维绘图总体来看仍是当前 CAD 的主导形式,例如日本工业界现在 97%,甚至 99% 的 CAD 工作量用二维绘图软件完成。二维模型的优点是简单、实用,对当前的生产管理

体制产生的冲击最小,它能表示任意复杂的产品结构。当然,产品越复杂,图面上的视图、局部剖视越多,图纸也就越难读懂。因此,在 CAD 系统设计中,必须注重二维绘图功能的开发,建立产品图的几何和非几何工程数据库,使之使用灵活、方便。

2 . 三维线框、曲面、实体模型

三维模型作为二维模型的延伸具有明显的优越性,代表了 CAD 技术的发展主流。波音飞机公司从 1988 年起用发动机舱试点,探讨了用 CATIA 软件进行三维结构设计的可行性。此后在 777 新机设计中全面展开三维实体技术的应用,截至 1991 年 6 月实际达到的水平是 50% 的零件(共约 13 万件)用三维实体设计到 50% 的细节,在设计阶段排除了 50% 的差错,省去了部分样机的制造。美国麦道飞机公司以 UG 软件来建立全机的电子样机,完全省去了实物模拟和试切工序。

三维模型中的线框、曲面、实体技术代表了几何造型方法由低到高的三个发展里程。国外凡是历史较久的 CAD 软件,像 CADAM, CATIA, CV, UG , Bravo3, Calma, Intergraph 都是兼有三种建模方法,而且总的趋势是将三种层次模型有机结合起来,三者可以统一灵活使用,三维与二维模型从属于一个体系,相互协调一致。后起的软件像 I-DEAS 中的 Geomod, Pro/ Engineer, Euclid 则是侧重实体,将拼合运算做得十分可靠,主要依靠拼合以及局部操作来构造复杂形体。实体造型软件的几何处理至今有两种途径:

(1)采用 Romulus 软件 美国 Ian Braid 和 Charles Lang 等 1975 年创建 Shape Data 公司,所开发的 Romulus 软件,采用精确表示的一次和二次解析曲面。Romulus 的第二代叫 Parasolid,被 UG 采纳,成为 UG 的有机组成部分。它的第三代叫 ACIS。ACIS 核心班子只有三人,即 Alan Grayer, Charles Lang 和 Ian Braid,他们精益求精地改进 Romulus 算法,但仍保留了一次和二次解析曲面的底层数学基础。ACIS 正是他们三人名字和 Solid 的字首。

(2)采用多面体模型 采用多面体模型的好处是拼合运算中只要处理单一的平面求交和裁剪,缺点是数据量大,多面体逼近模型对于数控加工和精密测量不利,为了弥补这一缺陷,I-DEAS/ GEOMOD 采用了双重表示,既能查询 NURBS 曲面的原始精确定义,又用多面体进行拼合。I-DEAS 正在努力转向单一的 NURBS 曲面实体模型,这有很大难度。

3 . 特征模型

特征模型被公认为几何模型发展的下一代,它在更高的层次上表达产品的功能和形状信息。人们不满足于利用现有的 CAD 系统进行低层次的纯几何元素调用,而希望针对所设计零件的功能和总体布局直接组装功能形素,并且在设计过程中利用形素所固有的属性作出自动推理和决策,减少人工干预。特征模型就是针对自身的应用特点形成新的数据结构和新的几何拓扑描述的方法,将在本章的 4.4 节介绍。

4 . 生物模型

日本的生产技术过去一直学美国,无人工厂和集中控制的 CIM 模式都不够灵活。目

前正在探讨下一步的发展道路,于 1987 年开展了生物模型生产系统的研究。生物模型生产系统要求每个生产环节有自发性、自律性和自相协调能力,出现问题就地解决。每个基层单位有自主权和主动性,但又照顾到整体,保证总体计划上相互协调一致。生物模型的基本单元称为 modelon,采用面向对象方法定义。

过去的几何模型系统要覆盖各种对象的应用,照顾面宽,系统算法复杂,管理困难。现在每个 modelon 各有自己的几何定义和操作方法,在软件实现上变成单一化。例如通常用的照相机要调焦距、调光圈、更换镜头,结构和操作复杂,而一次使用的胶卷相机与胶卷本身的性能相匹配,取消了可调参数。modelon 上下级与平级间通过 message board 交换消息,消息板与专家系统中的黑板机制相类似,只是取消了集中的控制机构。一个 modelon 启动后,激活相关的 modelon,通过消息板对话,谋求问题的解决。这里强调的是自发驱动(spontaneity)、自律决策 (autonomy)、分散管理 (decentralization)、自我组织 (self-organization)、自由结合(connection free)、可变化(transformability)、有柔性(flexibility)。

以室内布置为例:教室里有桌子、椅子、黑板等。桌子的选型由它自己根据教室的环境确定尺寸、形状,而椅子又与桌子相匹配。又如机车车厢的拖动,集中管理方式下要先统计整个列车的车厢载重量才能决定火车头的马力和速度,而在分散化的生物系统中每节车厢自主确定马力、速度,但保证后一节车厢跟上前一节,相互不脱钩,也不碰撞。根据以上思想,已在工业界建立了协作组,研究工作在逐步深入。

4 2 几何模型

4 2 1 几何模型的概念

计算机图形学特别对几何形体的计算机模型(几何模型)感兴趣,并着重研究如何将几何模型转换成可见的图形(图像)。需要注意的是,模型和图形是互相联系的两回事。在许多应用程序中,特别是计算机辅助设计(CAD)中,程序的 80% 涉及到实体的建模问题,仅仅 20% 用来生成它们的图形。大量应用把模型的图形表示作为实现其目的(分析、构造、数控等)的一种手段。而另外一些应用,如计算机艺术、计算机辅助制图、动画等则把图形看成其最终目的。

几何模型(geometric models)描述的是具有几何特性的实体,因而适合于用图形来表示。几何模型包括两个主要概念:拓扑元素(topological element)和几何元素(geometric element)。拓扑元素反映了形体的空间布置,包括点、边、环、面、实体等,构成一个层次结构,实体由二维空间的封闭面组成,面由一个或多个封闭环界确定,而每一个环又由一组相邻的边组成,边由两点确定,点是最基本的拓扑元素。一个几何模型的所有拓扑元素构成其拓扑结构信息,反映了物体的几何元素间的邻接关系。

几何元素则是指具有几何意义的点、线、面等,它们具有确定的位置(坐标)和度量值(长度、面积等),所有的几何元素构成了几何模型的几何信息。

4 2 2 几何模型的建模方法

用几何模型来表示实体的方法有以下几种：

1 . 参数形体及其调用

产生几何形体的一种最直接的方法是对已有形体作线性变换,例如对一个单位立方体作等比例或变比例变换,会产生新的平行六面体,这些平行六面体都可看作是对原始立方体的某种调用。基本几何形体如正方体、球体、圆柱体、棱锥体等,称之为基本体素,对这些基本体素作简单的比例变换,就可产生各种新的形体。从图 4 .1 来看,通过变换而得到的形体与原形体之间的拓扑性质没有改变,但几何性质发生了变化。同样通过多次等比或不等比变换后,可产生各种相应的形体,如果把构成形体的各种尺寸参数作为变量,通过在调用这些形体的过程中指定相应参数,从而产生用户需要的各种产品形体。

图 4 .1 图例引用

2 . 扫描表示法

扫描表示法(sweeping representation)是基于一个点、一根曲线、一个曲面或一个形体在沿某一路径运动后生成一个二维或三维的物体。这种表示方法的关键是要先定义一个 $i - 1$ 维(其中, $i = 1, 2, 3$)的变换对象。这种方法已被用于许多模型系统中,并被公认为是对在某一方向具有固定剖面产品形体模型的一种实用而有效的方法。使用扫描表示法来进行实体模型设计需要两个条件:其一是给出一个运动物;其二是指定移动轨迹。运动物可以是曲线、曲面或实体,轨迹是可定义的路径。扫描表示法有旋转扫描和平移扫描两种主要的基本类型。

(1) 旋转扫描 在旋转扫描时,运动物上的每一点均在通过该点且与旋转轴成正交的平面上作圆周运动。以轴与平面的交点为圆心,以该点到圆心距离为半径确定的圆周运动,用这种方法得到形体的表面是旋转面。当被旋转的不是一条曲线而是一个二维实心体时,旋转扫描后的结果一般也是三维的实体,图 4.2(a)和(b)是相应示意图。旋转扫描法只限于具有旋转对称性的实体。

图 4.2 旋转扫描法

图 4.3 平移扫描法

(2) 平移扫描 平移扫描是将一个扁平的物体沿着某个指定的方向平移一段距离,得到的相应形体,因此这种方法实际上只要指定相应物体的横剖面,指定平移方向和距离就可以了。图 4.3 是这种方法的示意图,更复杂的形体,也完全可以用这种方法来得到。但平移扫描的范围只限于具有平移对称性的一些实体。

一般来说,无论用哪一种扫描方法,只要原来的二维图是可以解析表示的,扫描后的形体也有相应的解析表示。

3. 单元分解法

任意一个物体都可以分解成一些比原来物体容易描述的子物体。如一只咖啡杯,我们可以先把手柄和杯体分开,这是很自然的而且在拓扑上也是合理的,分开的手柄和杯体比原有杯子容易描述。进一步,可再把杯子分解成圆盘形的底和一个圆桶,它们还可以进一步分解,手柄也可以分解成若干小段。这种分解一直可以进行到每一子物体满足预定的可描述性标准。这一过程称为单元分解法(cellular decomposition)。任一物体都可以表示成由它分解而得的一组单元的“总和”。使用单元分解法的理由是原有物体无法描述,而分解后的单元可以描述。分解的方法有很多,每一种都不是惟一的,但它们都没有二义性。

单元分解法可以根据基本单元的形状,又分为四叉树、八叉树等情况。四叉树和八叉树分别是用层次式的二维或三维空间子区域划分来代替大小相等、规则排列的二维或三维单元阵列的一种高效方法。以下分别进行介绍。

(1) 四叉树 二维形体使用四叉树描述,它的表示是基于对形体所在的外接正方形递归地等分成四个子正方形。这种分解过程所形成的一棵树不像二叉树每个节点都有两个子孙,它的每个节点有四个子孙,除非到了叶节点。其层次结构中的根结点表示一个包围整个形体的正方形,如果形体充满该正方形,则不再分割;反之,要分成四个大小相等的小正方形。对每一个这样的小正方形,如果正方形是满的(即它是 full 单元)或它是空

的(即它是 empty 单元)则不用再分割;否则,当它是部分满 (partial 单元)时需再分成四个更小的正方形。按如此规则一直分割到不再需要分割或达到规定的层次限制为止,这样的过程称为四叉树编码。四叉树的根节点是表示整个形体所占的正方形区域。其叶节点表示不需要再分解的区域,这种区域的大小和位置是与 2 的方次有关。从给定节点到根节点的递归分解的深度取决于该节点在四叉树中的层次,也取决于该节点所代表区域的大小。图 4.4 给出了一个四叉树构造实例。用四叉树表示形体的精度取决于形体的大小、形状特征及其边界的曲率。几何数值越大,精度越高,处理所需时间也越长,存储空间也大。

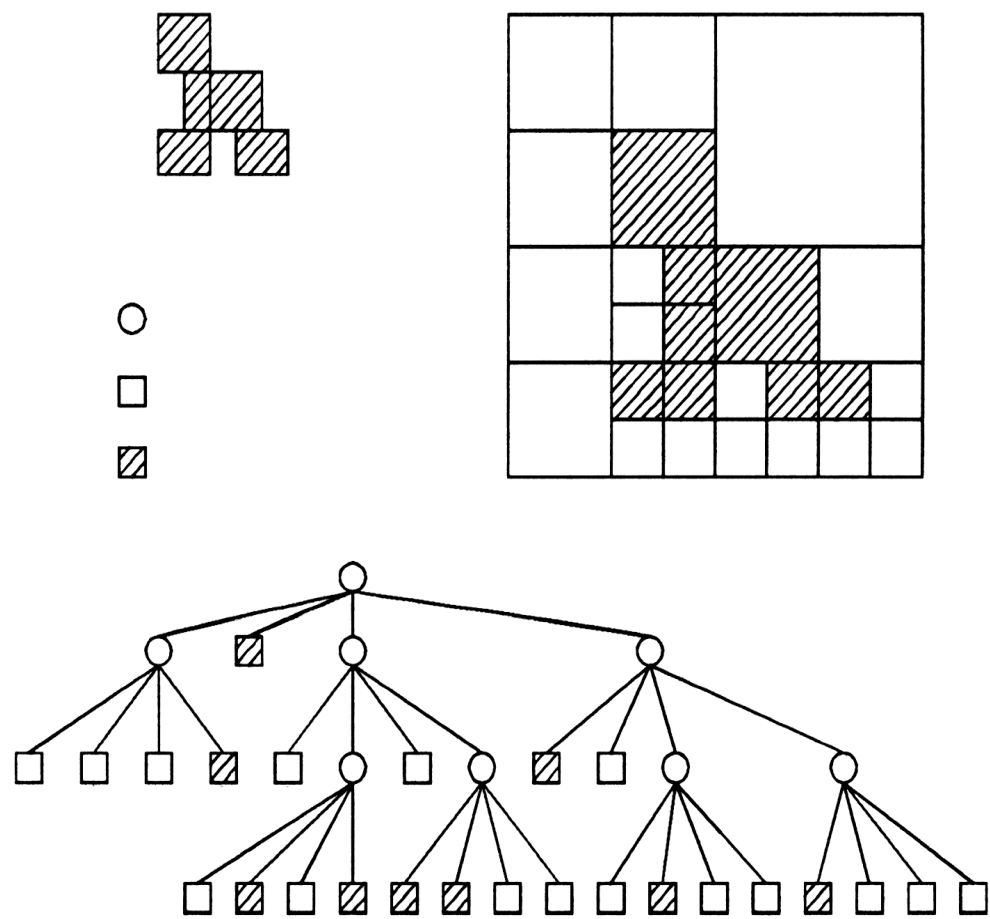


图 4.4 四叉树构造的实例

(2) 八叉树 八叉树是分层表示二维形状的四叉树编码的三维扩充。八叉树表示三维物体的原理是:先设定一个立方体的“物体空间”,要求其大到足以容纳我们希望表示的物体。然后,把该物体空间分成八个相等的正方体单元,称之为八分元(图 4.5)。在初始的八叉树结构中,有八个节点,每个节点表示一个八分元。每个单元在分割过程中有可能出现三种情况:全满(full)、部分满(partial full)、全空(empty)。对于全满或全空的单元不再分割;部分满的单元则继续分割为下一级的八分元。构成八叉树初始节点的一个子节点,直到最终生成的单元足够小,所有节点都不能再分割或达到一定精度要求为止。

4 . 几何体素构造法

几何体素构造法(constructive solid geometry, CSG),它是一种用简单几何形状构造复杂实体的模型方法,常被称为积木几何。CSG 的概念包括:规范化布尔运算符、体素、边界计算过程和点成员的分类。

图 4 5 用八叉树所示的实例

(1) 几何体素构造法的原理 几何体素构造法的物体形状的定义是以集合论为基础的,即集合本身的定义和集合之间的运算,也就是体素间的并、交、差运算。物体形状生成一个二叉树结构,树的节点是体素或变换参数,非叶节点是集合运算符并(union)、交(intersection)、差(difference)等。

由系统定义的简单形体,称为体素。现有模型系统为用户提供了一套形式简洁、数目有限的基本体素,这些体素的尺寸、形状、位置、方向由用户输入较少的参数值来确定。例如,大多数系统提供一种长方形体素,用户可输入长、宽、高和原始位置参数,系统可以检查这些参数的正确性和有效性。另外几种常用体素是圆柱体、球体、圆锥体和圆环。

较复杂的模型系统通过一些生成、删除操作也能产生类似的体素(如 GMsolid 和 Romulus 系统),系统也是通过检查输入参数来保证体素的确定性。体素的类型和数目的多少并不表示一个模型系统的功能强弱,例如,长方形体素和圆柱体素有相同的拼合和转换操作,可以构成由长方体、圆柱体、楔块、细长带状体、部分圆柱和四面体组成的体素系统。

CSG 模型系统除了最常见的拼合运算并、交、差三种形式外,也采用胶合(glue)算子,后者往往用于生成一些不规则的物体。

(2) CSG 表示法的实体描述 由于复杂物体可表示成为一些简单实体的某种组合,即通过并、交、差的组合操作。图 4 .6 表示了 CSG 构造实体的方法,可以清楚地看出实体的 CSG 表示是有序的二叉树,其中叶节点(或终结点)可以是体素,也可以是变换数据。非叶节点可以是规范化布尔运算(并、交、差),也可以是变换(平移或旋转)操作,它们都是用于下层的子节点(或子实体)。每个节点(不是变换叶)的子树表示由组合和变换操作所得的实体。根表示的是最终的物体。

5 . 边界表示法

边界表示法(boundary representation, B-rep)是以物体边界为基础的定义和描述几何形体的方法,并能给出完整和显式的界面描述。每个物体都由有限个面构成。每个面(平面或曲面)由有限条边围成的有限个封闭域定义,或者说,物体的边界是有限个单元面的并集,而每一个单元面也必须是有界的。

图 4.6 CSG 构造实体的描述

(1) 边界表示的原理和方法 在 B-rep 表示法中,一个实体通过包围它的曲面片或平面多边形(面片)的集合(边界)来表示。一个实体的边界必须将实体内部的点和外部的点区分开来,因此实体的边界完整地定义了该实体。用边界表示法描述实体,必须满足封闭、有向、不自交、有限和相连接,并能区分实体边界内、边界上的点。

边界表示法可以用一系列点和边有序地将实体边界划分成许多单元面。例如,实体的平行六面体可以方便地分成 6 个单元平面,各单元面由有向、有序的边组成,每条边则由两个点定义。圆柱体底和顶面自然也是一个单元面,而圆柱面的分割则有多种方法,图 4.6 中划分为前后两个圆柱面,每个圆柱面则由有向、有序的直线和圆弧线构成,而圆弧线则由三点定义圆的方法描述。不管是平行六面体还是圆柱体,都不能只用一个单元面表示整个物体,因此,每个单元面都是整个物体边界的子集。

(2) 边界表示法的数据结构 边界表示法强调物体的外表细节,建立了有效的数据结构,把面、边、顶点的信息分层记录,并建立了层与层之间的关系。分层记录的信息包括相互独立又相互联系的两部分:一组是几何信息,一组为拓扑信息。几何信息是指欧氏空间中的位置和大小,包括点(point)的坐标,曲线(curve)和曲面(surface)的数学方程等。拓扑信息是指几何体的顶点(vertex)、边(edge)、面(face)的数目、类型以及相互间的连通关系。根据这些明确的记录信息,可以知道几何体表面的范围及其邻接情况。

图 4.7 所示为一个棱锥四面体的边。根据边界表示法原理,它可以用一系列点和边有序地将其边界划分成许多单元面。因此,每个单元面都是整个物体边界的子集,或者说,一个物体所有单元面的并集,等于该物体的边界。但对于曲面实体,将需要更多的条件。

图 4.7 四棱锥的边界表示

图 4.7(a)的边界描述形式如图 4.7(b)所示,它可以看作是含有体、面、边、顶为节点的信息分层。这种数据结构关系建立了层与层之间的关系,一组是几何信息,一组是拓扑信息。因此,拼合过程的集合运算也分成几何信息处理和拓扑信息处理两部分。

在以上所述用几何模型表示实体的几类表示法中,使用较多的是扫描表示法、结构实体几何法和边界表示法三种。三视图重构三维实体常用于产品模型中,但上述五种方法中,没有一种方法能在各方面都比其它方法优越。因此有些系统常使用二种或更多种表示法,以对付不同的应用。这样就需要在不同表示法之间进行转换。

4.2.3 模型的操作

除了少数实体是整体的(不可分割的)以外,多数实体具有层次式结构。它可以通过一个由底向上的构造过程来产生。元件被用作构造模型以建立较高级的实体,所产生的实体本身又可能作为构造模型来建立更高级的实体,如此等等,因此实体的模型是一种多层次结构的模型。使用构造模型来建立更高级实体时需要对模型进行若干种操作:

1. 模型的变换

每一构造模型是在其自身的坐标系统中定义的。它必须变换到较高级物体的坐标系统中去。如果两坐标系统的方向一致,那么仅涉及平移变换;如果方向不同,则还要进行旋转变换;而当其形状相同而大小不同时,还要进行比例变换等。

为了建立更高级实体所需的对某些构造模型的使用称为图例调用。调用时所需的变换称为图例变换或模型变换。在一个层次式的模型构造过程中,在不同的层次都可能使用模型变换。对每一特定层的某个构造模型,有两类模型变换对它发生作用:一类称为局部变换,即直接作用于该模块的变换;另一类称为全局变换,即作用于由该模型参加构

造的高一层的模型的变换。在不同层次有不同的全局变换和局部变换定义。对于一个层的形体模型来说,它要经过一次局部变换和若干次全局变换才能成为最高层实体中的一部分,由于这和该形体模型经过由这一系列变换综合而成的复合变换的作用的效果是一样的,所以可在计算时先计算出复合变换矩阵,然后再对各个形体实施复合变换,这样做可以大大减少计算量。

2 . 模型的遍历

模型的遍历有三方面的目的:显示、修改和分析。为了遍历的需要,对模型中每一层次每一子模型必须进行编码处理,一般以整数序号作为每一子模型的标号。

由上可知,所谓模型实际上是有关的应用数据结构加上一些应用过程的集合,后者用来建立、修改、显示、分析该模型。

4 3 参数化模型

参数化(Parametric)设计是指参数化模型的尺寸用对应关系表示,而不需用确定的数值,变动一个参数值,将自动改变所有与它相关的尺寸,并遵循一定的约束关系。也就是说采用参数化模型,可通过调整参数来修改和控制几何形状,自动获得产品的精确模型。

近几年来,以 MIT 较早研究成果为基础,以 PTC 公司的 PRO/ ENGINEER 的参数化建模系统为先驱,参数化设计技术得到了迅猛的发展。SDRC 公司的 Variational Design和 MATRA 公司推出的 Adaptive Design 都是参数化技术迅速发展的产物。由此可见,参数化技术正日益完善并逐步走向商品化。

参数化模型有多种,如几何参数模型、力学参数模型等,本章主要介绍几何参数模型。

4 3 1 参数化设计

参数化设计的出发点是,通过说明产品的几何组成要素之间的几何特征及相互位置关系,定义一类形状或结构相似的几何实体,该类实体的实例由一组特定的几何约束值确定。因此,参数化实体模型是由几何模型和几何约束模型两部分组成。完备的约束模型通过尺寸对几何形状的某些控制元素加以约束,构成对几何形体数据的惟一而完整的表示。在产品的几何形状修改和重建时,可通过几何约束建立尺寸与几何定义变量之间的约束方程组,求解约束方程组以实现尺寸变化到几何改变的传递。

几何模型中,如果不包括几何约束信息,各元素间彼此孤立,不相互关联,则无法支持变参设计。因此,在参数化设计系统中,首先必须建立包括几何模型和几何约束模型的参数化模型。

在 CAD 系统的设计中,不同型号的产品模型往往只是尺寸不同而结构相同,映射到几何模型中,就是几何信息不同而拓扑信息相同。因此,参数化模型首先要体现零件的拓扑结构,从而保证设计过程中几何拓扑关系的一致。实际上,用户输入的草图中就隐含了拓扑元素间的关系。

产品零件的参数化模型是带有参数名的零件草图,由用户输入。几何信息的修改需要根据用户输入的约束参数来确定,由用户输入的参数(或间接计算得到的参数)的参数名找到对应的实体,进而根据参数值对该实体进行修改,实现参数化设计。图 4.8 所示为某零件参数化模型及各尺寸参数变量。

图 4.8 零件参数模型

图 4.9 产品参数化模型的构造

对于拓扑关系改变的产品零件,也可以用定义它的尺寸参数变量,来建立该零件的参数化模型。

参数化建模的表达模型应由 B-rep/ CSG 模式和约束模式共同组成一个有机体,使产品的参数化模型可以层次性地构造(图 4.9)。一般来说三维特征体素由二维形状采用“扫描表示法”生成;二维图形通过变动几何方法实现参数化,即先画出二维草图,然后定义尺寸及几何约束并指定“扫描”的类型;最终形成参数化的特征体素。参数化的体素在一定尺寸和连接关系下进行布尔运算,由此得到参数化的产品模型。

4.3.2 约束

约束(constraint)可以解释为若干个对象之间所希望的关系,也就是限制一个或多个对象满足一定的关系 $R(A,B,C,\dots)$,对约束的求解就是找出约束为真的对象的值。

根据产品设计意图,设计者要在几何模型中明确地表示出诸如尺寸等的约束条件,模型的尺寸反映了设计者的意图。约束主要应用在几何布置、仿真、支持设计、分析和推理等方面。约束的类型很多,以下介绍几种:

1. 几何约束

所谓几何约束就是要求几何元素之间满足的某种关系。利用几何约束,设计者就可以方便地表达其设计意图,不需考虑细节而能尽快地草拟零件图。

几何约束从性质上可分为结构约束与尺寸约束两类。结构约束指拓扑与结构上的约束,是不变动的对象,如对称、平行、垂直、相切等;尺寸约束则为固定几何元素间相对位置的约束,是变动的对象,如各种距离、两线的夹角、圆的半径等。几何约束可分为两类:第一类能够方便地转化为代数形式(代数约束),直接确定模型参数;第二类则需要经过专门的程序进行处理,由用户给出模型参数,再行验证。如约束 $\text{Inside}(A,B)$,表示 B 包含在 A 封闭多边形中。

2 . 数据一致性约束

数据一致性约束对保持几何模型的数据完整性十分重要。像图 4 .10 两线段 G_1 、 G_2 相交于 P_1 ,不但要描述它们之间的拓扑关系,而且还需说明 P_1 是 G_1 和 G_2 的交点,坐标关系中给出 P_1 点的坐标必须和交点坐标相同。

图 4 .10 两线相交于点 P_1

拓扑结构中,这样的约束条件就表示了模型的几何特征,尽管有时可能产生数据冗余,但一定要设法保证模型的一致性。

3 . 代数和非代数约束

一般情况下,几何约束在某种程度上都可以看成是几何元素的带参数的代数约束,如二线平行,用一阶谓词表示为 $\text{Parallel}(L_1, L_2)$,它可以写成带两个参数的代数约束:

$$\text{Parallel}(L_1, L_2) \quad \text{line-direction}(L_1, A_1, B_1) \quad \text{line-direction}(L_2, A_2, B_2) .$$

由于所有的几何元素都能根据其几何特征和参数化定义相联系,从而所有的几何约束都能看成为代数约束。如果是这样,那么所有的约束问题都可以从几何元素(公理性)级归纳到代数约束级(分析性)。当然,不是所有的几何约束都可用这些方法进行转换。

这种用带参数的代数约束来表示几何特征很不方便,像那些只存在某种几何关系,但几何元素间没有固定的约束关系的形体,必须用非代数约束来表示,一般采用逻辑约束表示之,如像非代数约束 $\text{inside}(LP_1, LP_2)$ 表示 LP_2 包在 LP_1 上,这可以通过几何计算来求值。非代数约束的另一个优点是:一些几何推理可以在公理级进行,而无需进入到代数级。例如直线 L_1 和 L_2 平行,而直线 L_2 和 L_3 平行,那么可以推出 L_1 和 L_3 平行。

4 . 功能约束

功能约束是对于功能的限制,并将这些限制综合成设计目标。如“盛水”,需要采用手持、放置等操作,则“手持”、“放置”功能的结合就是功能约束。如果功能映射为结构,功能

约束就可以转化为几何约束。

4 3 3 参数化建模技术

参数化建模的基本思想是用几何约束说明产品模型的形状特征,从而设计出一套形状或功能相似的设计方案。在此思想指导下,参数化建模技术记录建模的整个过程,修改尺寸即修改建模过程的某一数据。参数化模型方法按其约束表达模式不同大体可分为人工智能法、CSG/几何树表示法和几何法推理法三大类。

1. 人工智能方法

由 Sunde 等人提出的面向人工智能的方法,是用一阶逻辑谓词描述几何形体的约束关系,几何与拓扑结构用一阶逻辑谓词的形式描述,并写入知识库中;进而借助推理机确定图形元素(点、线、面)的关系进行建模。面向人工智能的参数化建模方法运用符号化的表达方式表示各种类型的数据,使系统对语义的表达更为直接。但由于目前还不能很好地处理约束一致性问题,以及存储空间消耗大,对数值计算和图形交互的支持能力弱等不足,因而该方法主要用于构造二维图形和简单的三维形体。

2. CSG/几何树表示法

CSG/几何树表示法是以 Emmerick 为代表提出的数据结构方法,它是在传统的几何模型信息上加入几何约束信息,采用 CSG 树和几何树构造三维形体,适用于构造拓扑关系发生改变的形体参数化。该方法的局限性在于:适用的几何约束范围小,且问题的解决需建立在成熟的算法基础上。

3. 几何推理法

几何推理法的原理是从已知元素和几何约束,通过推理逐步确定出未知元素。它与变分几何法相比,仅属于一种局部求解的方法。几何推理法又可分为两种:一种是建立在专家系统的基础上,采用谓词表示几何约束,通过推理机导出几何细节,这种方法可检查约束模型的有效性,并具有局部修改功能,但存在着速度慢、系统庞大、对循环约束情况无法求解等问题;另一种是扩展现有的图形数据结构,使其包含几何约束信息,并通过程序实现从几何约束到几何细节的推导。这一方法有希望克服以上方法的不足,是更为实用的方法。它应使系统不仅能处理简单的尺寸约束,而且能处理整体尺寸约束。

(1) 基于推理机的参数化建模方法 基于专家系统推理机进行推理的方法,其推理思想是:根据初始条件和提供的约束,经过推理,便可知道用哪些约束可以解出哪些“知识”,而根据这些“知识”,又可解出另一些,从而把可解的“知识”全部导出,直到获得最后解。可见几何推理是局部求解的,是一步步地推导出来的。其系统组成包括推理部分和几何计算部分两部分(图 4.11)。

推理部分将几何约束(尺寸、平行线、隐含约束等外部信息)转换成知识,并存入知识库,中间是一个推理机模块,和一个规则库(如图 4.11(a)所示)。所谓规则库,例如已知线的端点,且知其方向,可写出它的平面方程系数,把这些规则存入规则库中;又如已知圆

图 4 .11 基于推理机方法的系统结构

弧的两个端点和圆心,就可计算出该圆弧,也把这方程列出,存入规则库,然后把知识与规则进行匹配。例如知识库中有一个首点,且在某直线上,此线的方向也是已知的,如从规则库中搜索到有一条规则,就可推出其平面方程系数,从而求出未知元素。

知识库到规则库中去搜索,如搜索出来有此规则,就推理出来一个“知识”,而把它存入知识库中,然后又到规则库中去匹配。这样逐步地把几何元素推理出来。它的推理部分是一个专家系统,采用 LISP 等人工智能语言编程。而推理历程模块记录下了这个推理过程,然后再由几何计算部分来调用。

几何计算部分的系统结构框图如图 4 .11(b)所示,它包括构造计划及其解释器、执行程序、原型几何、精确几何模型等部分。在推理部分中,推理历程模块所记录的推理过程,就是一个构造计划,由构造计划模块取出后,将其内容送到构造计划解释器去进行解释。因为构造计划是一系列字符串,故需要构造计划解释器解释后,才可分辨出哪些是变量,需用哪些函数来计算,然后调用执行程序模块来执行,以得出精确几何模型。几何计算部分中的原型几何模块的功能,是用来克服多解的问题。因为有时一个方程可能有多个解,例如已知距离,而它的端点的确定就应根据草图所指定点的大致位置而定,也就是原型几何模块中存放了初始设计草图,从而避免了多解。所以整个几何计算过程为:从构造计划开始,经过构造计划解释器的解释,调用执行程序加以执行后,通过原型几何进行选择,最后得出精确的几何模型。

基于推理机的推理方法的几何计算部分常用高级语言(如 C, PASCAL 等)编程实现,若用 LISP 等人工智能语言编程,则效率将会很低。

(2) 基于几何推理的参数化建模方法 基于几何推理的参数化建模方法由三部分组成:第一部分是基于约束的参数化模型,该模型将几何约束蕴含于图形数据结构中加以表示,易于建立与应用;第二部分是通过程序实现几何推理的算法,该算法能有效处理整体性几何约束,并进行约束模型的一致性检查;第三部分是基于推理的参数变动处理算法,该算法具有局部修改功能,尤其能够同时适用于对当前设计图形及从数据库中调出的已有图形进行参数修改。图 4 .12 为其算法框图。

建立基于约束的参数化模型是实现参数化设计的前提。现有的几何模型主要由拓扑结构信息与几何信息两部分组成,前者反映物体的几何元素间的邻接关系,后者反映几何元素的位置与度量值。既然其不含有几何约束信息,因此无法支持参数化设计。建立参数化模型的关键在于如何有效地表示与建立几何约束。

图 4 .12 参数化算法框图

通过扩展现有的二维图形数据结构,将几何约束蕴含于图形数据结构中加以表示,建立起参数化模型。这种表示能够支持通过程序进行的几何推理,有利于提高建立与应用几何约束的效率。

对于结构约束,鉴于其表示几何元素之间的固定联系,不需标注与显示,信息量小,故我们将其直接蕴含在二维图形的点、边数据结构中。例如,表示相切约束,我们仅需在边的数据结构中设置相切边的指针;对于对称约束,在边的数据结构中采用对称边及中心线指针共同表示。

与结构约束相比,尺寸约束由于其可以变动,需要标注、显示,其所包含的信息量较大,所以可以专门设置一尺寸结构来表示尺寸约束。尺寸结构主要包含以下信息:尺寸类型,尺寸涉及的几何元素指针,用于进行尺寸标注、显示的信息,如尺寸的位置、名称、尺寸的度量值、尺寸的分解操作等。

为了提高几何推理与参数变动处理的效率,可在几何元素及尺寸的数据结构中专门设置若干信息,它们是已知元素标识符、已知约束标识符、几何元素与约束的次序号等。有关这些信息的意义与作用,后面将详细讨论。

对于几何约束的输入,可仍采用图形交互输入方式,在输入草图元素的同时,输入该元素相应的几何约束,并由系统自动建立起相应的数据结构。具体方法是对现有二维图形菜单中的几何元素,根据其可能有的约束类型进一步分细,例如,对于直线边,在新的图形菜单中被分成一般直线、水平线、垂线、中心线、切线等多种类型。既然所选的每一元素都有特定的约束语义,再配上约束参数,则在输入几何元素的同时,也输入了几何约束。

几何推理就是根据给定的几何约束确定出具体几何元素及精确几何模型的过程。具体内容包括两个方面:第一是从草图与尺寸约束推导出精确的几何模型;第二进行约束模型的有效性检查。

其算法的基本思想是:从已知的几何元素出发,根据给定的几何约束,推导出与该已知元素有约束关系的各元素的部分甚至全部信息,将无法求解的整体约束逐步转化成几何元素的局部约束,最后根据每一个几何元素的局部约束,调用相应的计算程序计算出每一个几何元素。约束模型的有效性检查则包含在几何推理过程中完成。在给出具体算法之前,先对数据结构中用于几何推理的两个信息单元加以说明。

在已知元素标识符时,我们称一个几何元素为已知,则该几何元素已经具备了完整、

精确的几何信息,否则称为未知。几何元素的数据结构中设置的已知性标识符,对已知与未知元素进行标识,以指导几何推理总是从已知到未知,提高推理效率。

在已知约束标识符时,称一个几何约束相对于某一约束元素为已知,则其另一约束元素也为已知元素。在几何元素的数据结构中,对该元素涉及的每一约束设置一标识符,对已知约束与未知约束进行标识,用以判定该元素是否已具备足够的局部约束进行求解,从而及时地将满足条件的未知元素转变为已知元素。

几何推理的具体算法分为三步:

第一步 建立已知元素表,最初的已知元素为用户给定的初值。

第二步 对每一已知元素进行几何推理,从该已知元素的数据结构中查找出所有与之有约束关系的元素,再对每一约束元素根据不同情况进行如下不同的处理:如果约束元素为已知,且对方元素中的相应约束为已知约束,则不作处理;如果约束元素为已知,但对方元素中的相应约束为未知约束,则提示用户该约束为多余约束;如果加上现有的已知约束,该约束元素便可求解,则调用相应的计算程序计算出该约束元素,并记录该元素的次序号,送入已知元素表;如果加上现有的已知约束,该约束元素尚不可解,则在其数据结构中记录该约束为已知约束;如果该约束为尺寸,再记录其次序号。

第三步 最终检查。遍历所有几何元素,如果还存在未知元素,则显示出未知的几何元素及它们已知的约束,提示用户缺少约束。

4. 参数驱动法

参数驱动法是基于对图形数据的操作和对几何约束的处理的一种参数化图形的方法,它是利用驱动树分析几何约束来对图形编程的方法。

(1) 参数驱动的定义 参数驱动机制是基于对图形数据的操作。以二维空间为例(对三维空间也适用):一个图形绘制完后,图形中的各个图形单元(如点、线、圆、圆弧和字符等,一般称它们为图形实体)全部映射到图形数据库中。不同的实体类型有不同的数据形式,其内容可分为两类:一类是实体属性数据,包括实体的颜色、线型、类型名和所在图层名等;另一类是实体的几何特征数据,对于圆有圆心、半径等,对于圆弧有圆心、半径以及起始角、终止角等。

由于参数化图形在变化时,不删、不增实体,也不修改实体的属性数据(这反映了原图的风格),因此完全可以通过修改原图形数据库中的几何数据来达到对图形进行参数化的目的。

(2) 判断要修改的几何数据以及修改方法 在二维图形数据库中,图形实体间的关系是不反映的(在三维模型系统中,用多表结构可以容易地找到这种关系),这就要求建立几何数据与参数的对应机制。该机制是通过尺寸标注线建立的。尺寸标注线可以认为是一个有向线段,即向量(图 4.13)。上面标注的内容就是参数名;它的方向反映了几何数据的变化趋势;它的长短反映了图形现有的约束值,即参数的现值;它的终点坐标就是要修改的几何数据。其终点称为该尺寸线的驱动点,驱动点的坐标可能存在于其它实体的几何数据中,称这些几何数据对应的点为被动点。

当我们给一个参数赋新值时,就可以根据尺寸线向量计算出新的终点坐标,并以此来

图 4 .13 有向线段

修改图形数据库中被动点的几何数据,使它们得到新的坐标和新的约束。

例如图 4 .14(a)中所示尺寸线 **d** 可以看作是由(0,0)到(2,0)的向量,其长度为 2,是参数 a 的值;方向为 0° (与 X 轴正向夹角),说明 B 点将沿水平方向变化;终点 D (与 B 重合)就是驱动点,其坐标(2,0)就是要被修改的几何数据。通过 D 点可以标识直线段 l 的一个端点 B , B 就是被动点,给参数 a 赋值为 5,可算出新的终点坐标(5,0),将它替换原数据库中驱动点、被动点的坐标,即几何数据(2,0),线段 l 就伸长,变成了 l ,尺寸线 **d** 也变成了 **d**,如图 4 .14(b)所示。

图 4 .14 几何数据的修改

上例中如果参数 a 的值仍赋 2,则终点不变,驱动点、被动点的坐标就都不必修改。可见参数值的变化是这一过程的原动力,因此称之为参数驱动机制。

尺寸标注线有线性尺寸、直径尺寸、半径尺寸、角度尺寸等,对应的驱动点应具有广义的含义。

(3) 约束联动 通过参数驱动机制,可以对图中所有的几何数据进行参数化修改,但只靠尺寸线终点来标识要修改的数据是不够的,还需要约束间关联性的驱动手段来约束联动。

就二维情况来说(对三维也适用),一个点有两个自由度,确定它就要有两个约束条件。如果采用参数驱动机制就要标注两个尺寸线,若该点的约束之间有某种关系,或与其它点的约束有关系,则只需一个约束或可由其它点来确定。对于一条线段,由两个点或一个点(在共点时一个点可视为两个点)、一个角度和一个距离来决定,共四个自由度,确定它就要四个约束。如果能确定它们之间的相关关系,就可以任意控制这条线段的变化:或旋转,或伸缩,或平移,或者两两复合变化,圆弧亦可如此。把这种通过约束的关系实现的驱动方法称为约束联动。

推而广之,对于一个图形,可能的约束十分复杂,而且数量极大。而实际由用户控制的即能够独立变化的参数一般只有几个,称之为主参数或主约束;其它约束可由图形结构特征确定或与主约束有确定关系,称它们为次约束。对主约束是不能简化的,对次约束的简化可以用图形特征联动和相关参数联动两种方式来实现。

所谓图形特征联动就是保证在图形拓扑关系不变的情况下,对次约束的驱动,亦即保

证连续、相切、垂直、平行等关系不变。反映到参数驱动过程就是要根据各种几何相关性准则去判识与被动点有上述拓扑关系的实体及其几何数据,在保证原关系不变的前提下,求出新的几何数据,称这些几何数据为被动点。这样被动点的约束就与驱动参数有了联系。依靠这一联系,使与被动点存在几何或数据约束联系的从动点得到了驱动点的驱动,驱动机制则扩大了其作用范围(域)。

例如图 4 .15(a)中 $BC \perp AB$, 驱动点 D 与被动点 B 重合。若无约束联动, 当 $s=5$ 时, 图形变成图(b)的形状。因为驱动只作用到 B 点, C 点不动, 原来 $AB \perp BC$ 的关系被破坏了。经过约束联动驱动后, C 点由于 $AB \perp BC$ 的约束关系成了随被动点 B 而移动的从动点, 它也将右移 3 个单位, 以保证原有的垂直关系不变。

图 4 .15 图形特征联动

所谓相关参数联动就是建立次约束与主约束在数值上和逻辑上的关系。在图 4 .16 中, 主参数有 s, t 和 v 。设 s 由 3 变为 5, 根据参数驱动及图形特征联动, 图形变成了图 4 .16(b)状态。原来的拓扑关系没有变, 但形状已经不正确了。

为保证形状始终有意义, 就要求 $v > s_0$ 。假设我们能确定 v 与 s 有一个确定关系: $v = s + 2$, 那么就要有一种办法能标识这样的关系, 并保证实现这种关系。

具体实现是将这个关系式写在尺寸线上, 替换原来的参数 v , 图 4 .16(c)所示。这样该尺寸线所对应的约束就是次约束, v 就成了 s 的相关参数。在参数驱动过程中, 除了完成主参数 s 的驱动外, 还要判断与 s 有关的相关参数, 并计算其值, 再用参数驱动机制完成该参数的驱动任务(如图 4 .16(d)所示)。

图 4 .16 相关参数联动

相关参数的联动方法使某些不能用拓扑关系判断的从动点与驱动点建立了联系。把相关参数的尺寸终点称为次驱动点, 对应的被动点和从动点称为次被动点和次从动点。于是可以得到一个驱动树(图 4 .17)。驱动树表示了一个主参数的驱动过程, 它的作用

域, 以及各被动点、从动点、次被动点和次从动点与主参数的关系; 同时也反映了这些点的约束情况。图中由驱动点到被动点, 由次驱动点到次被动点的粗箭头表示参数驱动机制; 由驱动点到次驱动点的虚箭头表示相关参数联动, 是多到多的关系 (其实是通过参数相关性建立的关系, 而不是由点之间建立的关系); 由被动点(次被动点)到从动点(次从动点)的细箭头表示图形特征联动。有时一个从动点(次从动点)可能通过图形特征联动找到其它与之有关的从动点, 因此图形特征联动是递归的, 驱动树也会有好几层。

图 4 .17 驱动树

从驱动点(次驱动点)到从动点(次从动点)是一个参数(不一定是主参数)的驱动路径。不同的主参数有不同的驱动树。不同的驱动树和驱动路径在结点上可能有重合点。重合的次驱动点(树间重合)表明相关参数与多个主参数有关系; 重合的被动点、从动点表明该点受到多个约束的控制。这样就可判断出约束的情况。

以生成图 4 .18(c)的图形为例, 对(a),(b)中的闭合多边形 P , 利用约束的图形特征联动, 需由四个参数 x, y, x_1, y_1 来定义(如图 4 .18(a)所示); 再经过相关参数联动, 只要主参数 x, y 两个参数(如图 4 .18(b)所示)。参数驱动树如图 4 .19。图中有两棵驱动树($x,$

图 4 .18 图形特征约束

图 4 .19 x, y 驱动树

y 驱动树)、四个驱动路径(x, x_1, y, y_1 路径)。不同驱动树间有重合的次驱动点“4”,表明该点受到两个约束;从动点“3”也有重合,“5”既作为被动点(y 驱动路径中)也作为从动点(x_1 驱动路径中),这都表明“3”,“5”点受到多个约束。但由于各点每次约束的方向不同,所以就不存在过约束问题。驱动树方法可以直观地判断图形的驱动与约束情况,是一种很好的分析手段。

从图 4 .18 中我们会发现(a),(b)图形只表示了图形特征约束,而没有反映相关参数约束,“ x ”的标注线比“ $x_1 = x - 1$ ”的短。这说明相关参数可以对图形进行重新定义,而且可以任意修改相关性,表示不同的设计思想。由此我们引出图形编程的参数化思想。

4.4 产品的特征模型及建模

特征建模是一种更完整地描述几何的实体模型技术。这种技术对几何形体的定义不仅限于名义形状的描述,还应包括规定的公差、表面处理以及其它制造信息和类似的几何处理。这种包含制造等信息的模型方法称为特征模型(feature modeling),这种面向设计过程、制造过程的特征模型方法,克服了几何模型的缺陷。因此特征建模被公认为几何模型发展的下一代,是集成制造和智能制造的一种理想的产品模型。

4.4.1 特征的定义

自 80 年代以来,许多学者都对特征进行了定义。C .Hayes 和 P .Wright 把特征定义为“被连续加工过程切除材料的形状”;S .C .Luby 和 J .K .Dixon 把特征定义为“一个几何实体,该实体和 CIMS 中的一个或多个功能相关”,Dixon 又指出:“特征是具有形状与功能的双层属性的实体”;J .Shah 提出:“特征是一个形状,对于这类形状,工程设计人员可附加一些工程信息特征、属性及可用于几何推理的知识”。可见,可以将特征定义为:由具有一定拓扑关系的一组几何元素构成的形状实体,它对应零件上的一个或多个功能,能够被固定的加工方式加工成形。

从构造零件信息模型的目的出发,从智能制造和并行工程来看,特征是一组与零件的描述相关的信息集合。这里所说的描述可以是出于设计或制造、检测、管理等目的,例如从制造的观点看,特征表示与制造相关的形状和技术属性,从形状设计观点看特征是由几何和拓扑元素组成。所以在不同环境下,存在着不同的信息集合,因此,可以说,特征可定

义为一定环境下某种功能的反映。

随着应用的不同,“特征”的定义亦各不相同,即使是同一“特征”也可能有多种描述形式。但从设计角度来看,特征是功能与结构的对应几何描述;从工艺规划的角度看,特征是结构与加工工艺的对应方法描述;从加工的角度看,特征是加工工艺与机床走刀的对应过程描述。基于特征的产品模型就是包含设计、工艺规划、数控加工信息等特征的组合。

在具体系统中,特征被定义为一种参数化的形状单元,它具有几何、属性、制造知识三方面的信息,它同时满足设计和制造的应用要求。

可见,特征是从设计和制造经验中抽象出来的形状单元,它与制造工艺和制造环境相关,因而它不是一个单纯的几何实体,这就是特征与传统模型系统中体素的区别。特征是参数化的实体,实体上包含了一组待加工的“型面”。特征是设计中“体素”的概念与加工中的“型面”的概念的综合。从制造的角度来看设计者的意图,可以认为设计者定义了零件的最终形状,而制造者则必须在实际加工之前定义出毛坯至零件最终形状的一个或若干个形状变化序列。所以说,基于特征的产品建模就是对产品零件的产品模型的特征的描述,是指通过计算机模型化处理,将工程图纸所表达的产品信息抽象为特征的有机集合,使特征作为产品定义的基本单元。该模型不仅能支持各工程应用活动所需的产品定义数据,而且能提供符合人们思维的高层次工程描述术语,并反映工程师的设计、制造意图。

当已知某一特征的几何、属性和制造知识时,可以按照下述步骤来设计这一特征。

选定一组控制几何形状和属性值的基本参数变量;

根据选定的几何控制参数,推导出其它形状参数的约束关系和计算方法,确定参数实体的拓扑结构的构造方法;

定义一组几何控制参数的合法校验规则和拓扑校验规则,分别用于检验输入参数和布尔运算的结果;

定义特征所具有的属性类型,如形状公差、位置公差和粗糙度等,属性记录的初始化方法,主-子特征间属性的继承关系和引用关系,属性值的校验规则;

定义各个待加工表面的成形方法,加工工艺片段,相关的数控子例程,装配关系;

定义特征的信息交换方式,即人-机界面的一个片段。

采用特征的概念所描述的产品模型是信息更加完善的模型,它的数据结构是集几何和属性为一体,是 CSG 和 B-rep 表示的结合。

4.4.2 特征的分类

特征可分为模型特征以及面向过程的特征。模型特征是指那些实际构造出零件的特征,而面向过程的特征并不实际参与零件几何形状的构造,它是那些与生产环境有关的特征。

模型特征进一步分为基本特征和二次特征。基本特征指构成零件主要形状的特征,设计用基本特征,以参数化形式存储在特征库中,这是通过特征的一些属性参数来表示整个数据实体的隐式表达法,它用最少的信息来定义形状特征,是一种十分简单明了的表达方法。由于它参数化地而不是几何地将形状信息组织进数据结构中,所以更易为后续应

用所控制和操作。在零件设计过程中,这些参数化特征可用实际特征值实例化处理后放在指定位置。二次特征则指用来修改基本特征形状的特征。二次特征又有正负之分,正特征用来描述如凸台、筋板等几何实体,负特征则描述孔槽之类的形体。对于每一正特征或负特征,其工艺参数如形状公差、尺寸公差、粗糙度等可由相应的属性加以描述。

更具体地说,由于零件特征是其设计制造等方面的信息,从零件的使用功能、制造方法等角度出发,大多数专家们都认为零件可用以下几种特征类型来描述:

(1) 形状特征 形状特征(form features)用于描述具有一定工程意义的几何形状信息。形状特征是产品信息模型中最主要的特征信息之一,它是其它非几何特征信息(精度特征、材料特征)的载体,非几何特征信息作为属性或约束附加在形状特征的组成要素上。形状特征的多级抽象表达,提供了表达精度、材料特征的实体,可以在分别表达精度、材料特征后,通过指针与形状特征联成网络。

形状特征又可分为主形状特征和辅形状特征。主形状特征(简称主特征)用于构造零件的总体形状结构;辅形状特征(简称辅特征)用于对主特征的局部修饰,它依附于主特征上。

(2) 精度特征 精度特征(precision features)用于描述几何形状和尺寸的许可变动量或误差。例如尺寸公差、几何公差(形位公差)、表面粗糙度等。精度特征又可细分为定位尺寸公差特征、定形尺寸公差特征、形状公差特征、位置公差特征、表面粗糙度等。

(3) 材料特征 材料特征(material features)用于描述材料的类型与性能及热处理等信息。如性能/规范(机械特性、物理特性、化学特性、导电特性等)和材料处理方式与条件(如整体热处理、表面热处理等)。

(4) 装配特征 装配特征(assembly features)用于表达零件在装配过程中需用的信息。

(5) 性能分析特征 性能分析特征(analysis features)用于表达零件在性能分析时所使用的信息,如有限元网络划分等,有时也称技术特征。

(6) 附加特征 附加特征(additional features)用于表达一些与上述特征无关的零件的其它信息,如用于描述零件设计的成组技术(GT)码等管理信息的特征,也可称之为管理特征。

4 4 3 特征建模

由于加工的环境、生产的规模、产品的相似性程度、标准化和系列化程度各异,针对各种情况,其建模方法也各不相同。

采用特征技术构造产品信息模型,以形状特征为主干,而且挂靠在此主干上的信息也都用特征来表达的零件模型,更符合工程实用和机械加工习惯,对这些信息的访问也非常方便和快捷。

1. 形状特征

可以将形状特征作为产品模型的主干,因为它是设计和制造过程中的通信媒介,具有特定功能,且与一定的加工方法相对应。无论从设计者角度,还是从制造者角度,都是从

形状特征开始处理问题,同时形状特征具有与零件其它属性很大的相关性。

(1) 形状特征定义 根据各形状特征在构造零件的几何形状、满足零件的功能要求和制造要求中所起作用的不同,将它分为主特征与辅特征。主特征是用于构造轴类零件的总体几何结构形状的特征,如圆柱面、圆锥面;辅特征是附在主特征或辅特征之上的特征,是对主特征或辅特征的进一步修饰。

根据形状特征之间的邻接关系(主特征与主特征之间的关系)和从属关系(主特征与辅特征之间的关系),可以作出一个零件(如图 4 .20(a)所示)的形状特征关系图(如图 4 .20(b)所示)。此多叉树结构的形状特征关系图,可以转化成易处理的二叉树结构(如图 4 .20(c)所示)。这种二叉树结构就是所要建立的零件信息模型主干。

图 4 .20 一个零件形状特征关系图及二叉树

如将产品零件各主辅特征的有关数据挂在产品信息模型的主干上,再辅以其功能要求信息、产品管理信息(包括标题栏信息、GT 码等)和技术信息(包括未标注表面粗糙度、总体热处理要求等),以及轴向尺寸信息,就形成了轴类零件信息模型的总体结构(图 4 .21)。

在实体模型的研究领域里,人们已经研究过多种形状定义的方法。但是,由于受到传统模型系统设计思想的束缚,许多种形状定义的方法都是为设计某个具体的形状和体

图 4 21 零件信息模型总体结构

素而提出的,没有从设计与制造的形状定义的一致性方面来考虑工具化系统的形状定义问题,或称特征的参数化几何实体的定义问题。例如在第一代的模型系统中就采用了“平行扫”的形状定义方法,它所定义的是一个具体形状,用户无权使之变为一个参数化的实体。这里提供的形状定义方法,是一种定义参数化实体的方法。

目前已经采用的形状定义法则有平行扫法则、回转扫法则、箱体法则、曲面立体法则和异形体法则等五种,此外还可进一步扩充一些新的法则。

平行扫法则是定义一个 2D 截面形状,沿 3D 直线轨迹扫动,如图 4 3 所示,这样可以构造出 2.5D 的实体,其表面可以是平面或柱状面。所有的 2.5D 实体都可以用这种法则来构造。

回转扫法则是定义一个 2D 截面形状,绕 3D 直线轴回转扫,如图 4 2 所示,这样可以构造出回转体。传统模型系统中所采用的圆锥、圆柱、圆球等体素就是回转体中的一个子集。

箱体法则是定义一个 2D 截面形状,沿一个封闭的平面导线扫动,如图 4 22 所示,这样可以构造出各种复杂的箱体。一般箱体都是由一个箱的主体加上一些凸台、孔、筋板等

图 4 22 箱体

结构构成,且主体上带有一些细微的圆角结构。箱体法则适合于构造箱的主体部分。若导线改为空间曲线还可以进一步扩大模型能力。

曲面立体法则是定义一组截面,经 B 样条插值处理后构成双三次 B 样条曲面(图 4.23)。在用户界面上可提供多种灵活的曲线、曲面定义方式,可以满足各种复杂曲面壳、叶片、杆、梁形状定义的需要。

图 4.23 曲面立体

异形体(如图 4.24 所示)法则,这种定义方法与曲面立体法则相似,其不同之处在于依据一组给定的截面形状只作线性插值。采用该法则可定义各种变截面实体,各种棱锥类杆、梁、筋板结构。

图 4.24 异形体

上述五种建模方法统称为“广义扫描模型法”,它们具有下述一些主要的特点:它们所定义的每个形体上的每一表面都与某一截面元素和一导线元素具有对应关系,称这种可以用来惟一地标识某一体素上某一表面的三元组(体素号、发生线元素号、导线元素号)为 CSG-索引。利用 CSG 索引可以方便地定义每个表面的属性;在 B-rep 模型上的每一个面记录中都含有 CSG-索引,从而通过 CSG-索引将 B-rep 模型与 CSG-模型连接在一起。广义扫描模型方法具有统一的形状定义形式,二维变异式设计方法可以经广义扫描推广至三维形体定义,同时它具有造型功能强和易于编程的特点。

(2) 形状特征二叉树 零件的形状特征二叉树各节点内容如图 4.25 所示。它的特征主属性表用于描述形状特征主要属性,包括特征号、特征类型、热处理要求、表面粗糙度值等。此外,还有一个较特殊的属性,即功能属性,用来描述该特征的功能,并由零件的总体功能进行功能分解而得。特征的功能决定特征类型,如轴类零件各形状特征的功能主要有:传动、定位、联结、支承密封等。

形状特征二叉树的几何尺寸包括定形尺寸和定位尺寸。定形尺寸是指描述形状特征几何形状的参数,如圆柱面的直径 D 和长度 L ,平键键槽的长 l 、宽 w 、深 d ,特征的各定形尺寸及其属性用双向链表串接起来:定位尺寸分轴向定位尺寸、径向定位尺寸、周向

图 4 25 形状特征二叉树各节点的内容

定位尺寸和方位角度。形状特征二叉树的形位公差链表用于描述产品零件的形状公差和位置公差。各形位公差及其属性用双向链表联接,形状公差基准一律为 NULL,而位置公差的基准至少有一不为 NULL,公差基准用特征号表示。

这种基于形状特征二叉树的产品模型清晰、简洁地反映了产品各特征表面之间的关系。在构造该模型时,抽取了产品信息的共性,即产品的形状特征二叉树结构及挂靠其上的信息类型,同时也充分考虑到零件信息的个性,即对不同零件,其形状特征二叉树一般不同,附于其上的信息一般也不同。对于个性,可用指针和链表结构来描述,各项内容根据具体零件或取值或赋空。这种动态的数据结构所表达的产品信息内容灵活,便于修改,还避免了构造繁琐的特征类库,节约了计算机资源,比较适合以设计者为中心的并行设计。

该模型中形状特征采用隐式表达,完全摆脱了几何模型中点、线、面的约束,但形状特征的隐式表达可以通过计算得到它的显示表达,这为特征模型和实体模型之间的联系提供了一条自然途径。

产品功能要求及特征中功能属性的引入,为建立产品功能与产品特征之间的映射联系创造了条件,便于产品设计与工艺设计中的各知识源,在满足产品功能的前提下相互协同作业,设计出最优的零件。

2. 属性特征

在制造和检验过程中两类最重要的属性信息是几何公差和粗糙度。几何公差包括尺寸公差、形状公差和位置公差,其中尺寸公差和位置公差都与基准体系有关。在机械工程图中,位置公差的基准是显式定义的,而尺寸公差的基准则是隐含的。这种隐含的基准在通用机床的制造环境下分为不完全重合的两个基准体系;其一是设计基准体系;其二是工艺基准体系。采用基于特征的模型方法必须统一设计基准与工艺基准。在指定形位公差时通常还伴有附加约束条件,如最大实体条件(MMC)、最小实体条件(LMC)、延伸长度、误差变化趋势等。依据 ISO 标准定义,形状公差分为:直线度、平面度、圆度、圆柱度、线轮廓度和面轮廓度六种;位置公差分为:平行度、垂直度、倾斜度、同轴度、对称度、圆跳动、全跳动、位置度等八种。尺寸及尺寸公差通常是工艺过程规划的主要线索,而形位公差将有助于正确地选择装夹定位方式。

在工艺规划时必须考虑的两个重要因素是加工表面的粗糙度和硬度。硬度值通常是采用磨削加工的依据,而粗糙度值是选择加工步骤及切削参数的依据。粗糙度的主要内容可以概括为成形方式(去除材料、不去除材料、其它)、表面参数和值及附加检验参数(采样长度、评定长度、平均间距和加工纹理及加工方法)。几何模型的空间尺寸关系可简化为两种形式:其一是反映两几何要素间的距离关系;其二是反映两几何要素间的角度关系。通常用于描述一个结构或一个特征的尺寸又可分为形状定义尺寸(如长、宽、高、半径等)和位置定义尺寸(如中心线、原点、圆心等)。

一般而论,形状尺寸只与一个体素上的一个或多个元素面有关,而位置尺寸则可能与多个体素上的多个元素面有关。如前所述,CSG 是一个精炼的概念模型,是一种向量形式的数字模型,具有可参数化的特点,其布尔运算树是一种层次状的语言表示。B-rep 是一种边界描述,具有详细的边界信息和拓扑信息,易于显示和人机交互。CSG 模型应与 B-rep 模型连接起来,以支持参数化和交互式两种方式的模型定义。采用 CSG-索引可以把 B-rep 模型与 CSG 模型链接起来。如果 B-rep 模型是一个离散化的多面体模型,则它的表面与原始 CSG 模型上的表面可能是一对一的关系,如果 B-rep 模型中的表面经 CSG-索引指向 CSG 模型中对应的体素,而 CSG 模型中体素、元素记录又分别与产品信息表相连,则经过 CSG 模型(参数化模型)和 B-rep 模型(交互式模型)都可以惟一地定义一个加工面的属性记录。

3. 特征建模系统设计

图 4.26 示出了特征建模系统设计包含下述组成部分:用户界面管理系统、特征库管理系统、布尔运算处理器、几何与属性的一致性校验系统、特征库、设计与制造知识库、工程数据库等。

特征建模设计系统的特征库用于特征的定义形式与其应用目的和环境密切相关。根据产品的不同,如机械类,以成组技术(GT)对产品零件作分类,依零件族进行特征分类,建立面向零件族、面向产品的特征库。

在设计阶段,产品零件按照其功能和几何的相似性被分为若干类型,如轴类、盘类、箱

图 4 26 特征建模设计系统组成

体类、支架类等。每一个特征库用于表达一类产品零件,用户可以按照自己的工艺方案、生产环境、产品对象来设计特征库。例如,根据产品零件的分类分析每一类产品零件的特点,从而抽取一组特征,这一组特征就可以构成该类产品零件的特征库。

特征建模设计系统中的知识库存放的是关于某一特征库中各特征可以共享的知识,而各个特征的加工知识则存放在各个特征的定义之中。例如常见的箱体零件可能有四种数控子程序,如平面加工、孔加工、轮廓加工、攻丝等;尽管构造各种箱体的特征可能有多个,而它们生成数控指令的方法可能是上述四种方式的组合。特征建模设计系统中的数据库用于存放工程设计和制造的有关数据。

特征建模设计系统中的布尔运算处理器,具有“交、并、差、装配”等布尔运算功能,其计算精度可以满足设计和制造的精度要求,其运算结果中各个交线边、环均带有凸凹性标志,由结果体的 B-rep 模型可以访问其对应的 CSG 模型,两者具有确定的对应关系。

特征建模设计系统中的一致性校验系统是面向特征的模型工具系统所特有的一个组成部分,它带有进行拓扑推理和符号推理的推理机(RM)。它可以根据某一特征的拓扑和属性校验规则对该特征在参与布尔运算后的性态进行校验,从而保证该特征的几何与属性的一致性。但有两种矛盾的现象存在:其一是,某一特征的几何与其属性相矛盾,例如某一柱面带有“孔”的属性,然而由于定位参数有误,使得实际的布尔运算结果并非是“孔”而是“槽”;其二是,某一特征的某一种属性与另一种属性相一致,例如,精度要求必须与相应的粗糙度要求一致,前者需用拓扑推理来校验,后者需用符号推理来校验。

特征建模设计系统中的特征库管理系统可以用于定义特征,如定义特征的参数化实体、属性信息模式、加工的知识规则、几何与属性的一致性检验规则等。它可以按成组技术(GT)的零件分族方法将特征分成与零件族对应的特征库,并实施库管理。

用户界面管理系统是一个用于设计人-机交互界面(如命令语言、菜单布置等)的工具。

4 5 智能制造环境下的产品模型

产品模型的建立达到了制造过程各方面对产品信息的共享,至于这些信息的产生、分析、修改和利用,部分地或全部地由人来完成。模型虽然提供了定义产品的全部信息,但许多信息是隐含的。这些信息对于无智能的计算机来说是无法理解、无法提供的。

4 5 1 智能制造对产品模型的要求

特征所具有的语义性与复杂性,要求用知识处理技术对特征进行处理。同时在特征的研究中出现的一些问题,如特征提取、特征映射等,也要求用人工智能技术解决之。在过去的研究中,Yoshikawa一直在进行人工智能技术在制造过程的应用,从一般设计理论出发,把产品建模和智能设计相结合,计算机通过描述有关产品的设计意图(design intent),在设计过程中定义功能特征。这样设计的结果就包括了符合设计者意图的各种功能特征。零件的技术信息,如公差、表面粗糙度等利用知识库里的规则,通过对功能与几何信息进行处理而得到。无疑地,在未来的智能制造环境下,产品建模应该和设计过程、工艺设计过程相结合,使人与计算机协同进行创造性的设计与工艺设计活动。这是制造活动本身的要求,也是未来发展的趋势。

开发一种适合于人与计算机协同进行创造性的设计与工艺设计活动的智能环境,是一个逐渐发展与完善的过程,这种发展是与人工智能技术的研究进展相关的。事实上,Yoshikawa等在开发实际的智能设计与产品建模系统方面所取得的进展是有限的,但他们提出在产品建模过程中,描述设计意图以及反映这种设计意图的功能与几何信息,在研究适合于智能制造环境的产品模型中是很有价值的。

1 . 适应智能制造环境的产品模型

综上所述,在研究与开发适应于智能制造环境的产品建模系统中,面向制造的产品模型的描述为:首先是描述产品零件形状的几何模型(实体模),它仅提供零件的几何信息;然后是描述产品的特征模型,它可以完整地描述产品的所有信息,可以作为产品数据交换模型,为制造过程的各方面所共享;第三是能够为计算机理解、提取产品信息,为制造过程的各方面所共享的产品模型,即为适应于智能制造环境的产品模型。

而面向制造的产品模型的表达应基于以下几个方面:首先是产品模型必须完整,模型的表达及所描述信息,应能适应知识处理的要求,提供面向各种应用的特征,并以反映不同语义信息的形状元素作为描述零件的基本元素,并用它们之间的约束关系描述它们的相对位置关系及所具有的性质,以适应在知识化处理的环境下,提供面向各种应用的特征;其次是计算机能够检查所表达的产品信息的一致性,避免二义性,计算机能正确地解释产品信息,探讨产品表达的形式化模式,使产品模型的表达和建模的方法建立在形式化理论基础上;第三是产品零件形状的实体模型,由描述零件的形状元素及它们之间的约束关系操作,用户和描述零件的形状元素打交道,而不直接操作实体模型。这样,描述零件的模型可主要划分为描述语义信息的模型和描述公称形状;然后是产品的装配模型通过零件的装配特征定义,由构成产品的零件、子装配体以及零件之间的连接关系等组成。

2 . 适应智能制造环境的产品建模

适应智能制造环境的产品建模,应该和产品的的设计过程相结合,在设计制造知识(包括设计手册中的数据、可制造性、可装配性、可维修性等方面的知识)和制造环境描述

(如加工各种特征的机床描述)的支持下进行,这样的—个产品建模系统包括了目前在CAD/ CAM 研究方面的所有内容。产品建模的结果不仅包含了产品模型本身,而且还包含了制造、检测、装配等方面的所有规划,以及面向管理的成本预算等信息,这样的建模系统是制造系统本身的要求。在这样的建模系统支持下,产品在设计过程中,考虑了制造、检测、装配等各方面的因素,对可制造性、可检测性、可装配性、可维护性进行评估,并对所描述的产品进行成本预算。在进行可行性评估中,自然要求进行相应的规划,这种思想也反映在“为了制造的设计”和“为了装配的设计”(DFM/ DFA)两个方面。事实上,在传统制造系统中,为生产做准备的技术部门之间的相互协作也体现了这种思想。

这种建模方式是人们追求的目标,计算机和人工智能技术的发展为有效地实现这种目标提供了可能。近几年,朝着这个目标建模的研究是制造系统领域研究的一个重要方向,如有人提出了以人与计算机协同进行产品建模的框架结构,有人提出了一个基于设计意图的有效地进行产品建模的系统框架。由于这种建模方法涉及制造过程的许多方面,因此有效地获取与之有关的设计与制造知识需要一个逐渐综合和积累的过程,同时对制造环境的描述也是针对特定制造系统的。利用这样的方式,开发有效地进行产品建模的系统,将是一个逐渐形成的过程。在开发中,应首先建立一个由人与计算机协同进行设计与制造规划的智能环境,通过对知识的积累(包括数量的增加和质量的改进),逐渐增强计算机的作用,从而把人的精力逐渐放到更高级的创造性活动上。

开发有效的产品建模系统的方式可分为三种:首先要与设计过程相结合,把产品建模过程作为一个设计过程,进行系统开发;其次是为了辅助设计与制造中的某些活动,要求有一个计算机表达的产品模型而开发产品建模系统;第三是为了进行智能制造的研究,为设计之后的制造活动提供一个产品表达模型而开发产品建模系统。第一种产品建模过程,一般先建立产品装配描述,后建立零件描述,或者产品装配描述与零件描述并行产生。第二种和第三种产品建模过程,一般是先建立零件模型,后建立产品装配描述。两者的区别在于完整、—致、无二义性地表达产品模型的程度,计算机智能地处理产品信息的程度。

4 5 2 产品的知识模型

1. 产品的知识

由于产品模型的发展要求计算机能够理解产品模型里关于产品的定义,并提供各种应用所需的高层信息。为此,需要把产品模型里的信息纳入知识的范畴,把人工智能技术应用到产品模型的信息产生、管理和利用中。由于产品模型里的信息由三维空间中产品的几何信息和附加在其上的属性组成,故把智能制造环境下的产品模型,称为产品知识模型。

知识模型的研究主要集中在如何将二维图形以一维的形式在计算机内表达和如何利用这些二维图形所描述的客观世界。可以借用这个概念来研究三维空间中的产品在计算机内的一维定义,以及将这种计算机内表达的对产品的定义以二维形式来描述,并在制造过程各阶段中应用。目的是建立在智能制造环境下的产品模型。

如前所述,约束反映了若干对象之间的关系,约束求解就是找出约束为真的对象的

值,这一特性特别适用于表达和推理产品的可视知识。制造过程是由许多不同阶段组成的,它们分别或一起影响着产品的成本、质量和整个制造过程的效率。制造过程各方面的这种相互制约关系形成一个复杂的关系网络,某一方面的变化都影响着其它许多方面。在现代制造系统中,由于计算机承担了制造过程中信息的表示和处理以及对过程的控制,因此有可能使产品在进行生产前就对其成本、质量和生产率进行评估和最优化。产品设计过程就是在满足制造过程中各方面要求下的信息处理过程,因此它可被抽象为满足一定约束的求解过程,基于约束的系统可适应对产品的多次修改和对设计过程的知识化处理,同样在工艺规划和生产过程中也有约束表示和求解的问题。

在产品描述过程中,产生与特征有关的知识,例如有了装配关系,就可以推出装配应用中的特征连接元素及其相关特性。与产品特征有关的知识都作为产品知识的一部分(它们是说明和附属在产品的几何形状上的),这样与产品知识模型打交道的任何一个部分都会面向它的特征。所以,特征就是产品知识模型和外界交换的信息。

特征是面向应用的,从产品知识模型里提取出特征,就必须要有推理功能,几何推理正是适应了这种要求。通过适当的推理规则确定或推断出形状的几何特性、参数及附加在形状上的有关属性就是几何推理。从基于约束的几何推理就可获取产品的知识,并对产品知识尤其是几何知识的修改以推断出面向应用的特征。

2 . 模型框架

产品的知识描述了在智能制造环境下的产品模型;约束用来表示和说明产品知识之间的关系(当然,在整个制造过程中约束的作用绝非仅限于此);特征提取描述了产品知识模型和外界的信息交换,对于形状获取而言,适合于人对形状的描述信息就是面向形状获取的特征,对工艺规划有意义的信息就是面向工艺规划的特征,以及对基于约束的几何推理来实现产品知识的获取、学习和利用。

在智能制造环境下,制造系统获取用户对产品的要求后,对产品进行功能设计、概念设计和详细设计的过程就是对产品知识进行自学习的过程。在这个过程中,系统对产品进行分析、优化和评估,产生可行的对产品的定义与说明。然后利用这个结果,提供产品的生产规划、工艺规划和加工指令的产生等过程所需信息。这就是对产品知识的推理和利用(图 4 .27)。

图 4 .27 产品知识的推理和利用

显然,从获取用户对产品的要求,通过自学习,完善产品的知识,得到对产品的定义与说明,在现阶段是很难实现的,除非产品的覆盖域限制在极窄的范围内。考虑到目前这方面技术发展的水平,以实现一个实用的系统,同时也考虑到技术的进一步发展对现有系统的再开发性,其策略是先建立产品知识的表达模型,然后直接获取其零件几何形状和它们

的装配关系等知识,知识的修改,自动或半自动地在人的控制下进行。在知识的利用方面,通过几何推理在面向应用的解释知识库支持下,产生面向应用的产品特征,实现对产品知识的共享。

4 5 3 产品的集成表示模型

制造过程的描述与建模是实现智能制造的一个关键环节。从前面介绍的概念可知,制造过程描述的生命期工程方法就是认为制造过程是由一系列对象所组成的,如产品过程、后勤、软件和制造者等诸因素,参与制造过程的每一个对象都是特定的、相关联的,并具有一定的有效期。为了适应智能制造系统中高度集成化与智能化的要求,通过集成知识工程、特征建模策略和面向对象的技术,建立一种基于知识的产品集成表示模型,以便为产品生命期中制造知识的处理提供一种框架。

在智能制造环境下,产品的集成表示内容应包括三个方面:数据、几何和知识。产品数据、几何和知识分别被定义为产品生命期内所有阶段附加在产品上的数据、几何和知识总和。数据包括公差数据、结构数据、功能数据和性能数据;几何包括几何图形、形状拓扑关系;知识包括特征知识和管理知识。因此,基于知识的产品集成表示模型如图 4 28 所示。该模型由若干个子模型互联而成,分属几何、数据和知识三种深度。从对产品描述的知识深度角度来看,自上而下深度增加,而抽象程度减小,各种深度上的每个子模型着重反映产品在该深度上的最小冗余度,使各子模型相互补充地形成一个完整的产品多知识深度表示模型。下面说明各个子模型的内容。

图 4 28 基于知识的产品集成表示模型

几何模型子模块是产品表示中最成熟和最基本的一个模型,由包括几何元素(坐标、点、线、面、方向)的多种定义形式来构成。拓扑模型子模块包含对产品的拓扑实体及其关系的定义,如顶点、边、面、路径等。目前常用的边界表示法(B-rep)可以较好地获取产品的拓扑信息。形状模型子模块是产品几何关系的数学表示,以几何模型和拓扑模型为基础,目前常用的表示方法是实体建模,即通过预先定义的一些体素,将产品表示成由这些体素构成的树结构或有向非闭环图,而体素的表示和各体素间的关系可分别从几何模型和拓扑模型中获得。结构模型子模块中的结构定义为一组具有语义的几何实体的集合,

包括一组几何实体及其相互关系和几何实体的语义表示两方面的内容。目前常用的方法是结构特征建模,该模型是公差模型和功能模型的基础。公差模型子模块反映产品中具有可变动范围的一类信息,它们是产品加工过程中一种重要的非几何信息,包括公差、形位公差、表面粗糙度、材料信息。功能模型子模块实际上是对结构模型中几何实体及其关系的语义各种功能的解释,可采用知识工程中的语义网络或框架来表示。一个产品的设计过程实际是从功能模型到结构模型的转化过程,因此,产品设计工作结束后,它的功能模型也就相应确定。性能模型子模块实际上是对产品的功能或结构按用户要求或预期进行的一种评价,主要包括性能参数、行为 (behavior) 值等,该模型与结构模型和功能模型是产品的可靠性设计和可维护性设计中的重要基础模型,它们将有助于解决目前复杂系统的监视与故障诊断领域中深层知识(如结构、功能与行为知识)的“瓶颈”问题。特征模型子模块包括产品几何特征和功能特征的参数化与陈述性描述产品生命期内各环节对产品结构施加的约束,它可采用知识工程中的知识表示技术。管理模型子模块的模型是对产品集成表示模型内部层次结构的描述,各子模型之间的关系、信息转换等,它可采用知识工程中的知识表示技术。

生命期工程方法,通过描述各对象在制造过程中各阶段的状态变化对其它对象的影响、性能评价、进入和离开该制造阶段的时间以及后期效应,来实现对整个模型在更高层次上的描述,它是一种颇有前途的研究方法。因为它将一个复杂的制造过程抽象为五种特定对象的生命期及其相互作用的过程,单个对象的生命期的描述工作是整个制造过程描述与建模的基础,现已有一些研究成果。基于知识的产品集成表示模型即是面向一个生命期的,该模型集成产品在生命期中各阶段的信息和对产品进行操作(如设计、加工、检验、维护等)所需要的各种知识源。然而,对不同对象的生命期之间存在的相互影响的进一步认识与详细描述,则还有待进一步研究。

第 5 章 智能 CAD 系统 及其设计

产品的设计活动是非常复杂的智能行为,涉及大量的设计领域知识和专家经验。为了加强和拓宽传统的基于图形学技术的 CAD 系统的功能,人们已经开始将人工智能和知识处理的技术用到产品设计领域,称为智能 CAD。人们用智能 CAD 系统来辅助完成设计方案的生成、选择和评价,以减轻其在设计活动中的劳动强度,提高产品设计质量和效益。本章就智能 CAD 系统的设计模型、设计方法和技术做一介绍,并介绍几个简单的智能 CAD 系统。

5.1 智能 CAD 系统的设计模型

5.1.1 设计模型简介

设计是一个面向目标的有约束的决策、探索和学习的活动,它根据设计说明给出对设计对象的期望功能(或行为)的描述,产生出符合设计要求的设计结果。在 CAD 系统设计中,设计模型(design model)的建立可以针对某个领域,也可以是通用的,而其目的是为实现具体的智能 CAD 系统提供理论依据。

目前,国内外专家提出的设计模型较多。除了基于 CASE 的设计模型将在下一节详细介绍外,这节将重点介绍国内外几种在 CAD 系统的设计中采用较多的设计模型。

1. 分析-综合-评价模型

分析-综合-评价模型(ASE)是由 W .Asimow 于 1962 年提出的,其主要观点是把每一个设计活动分解成为三个阶段(图 5.1)。图中的所谓分析就是对设计的理解问题,而且要形成一个对目标的显式的描述;而所谓综合是寻找可能的解答,通常可以通过目标分解法以及元素重组法来解决“综合”;评价就是确定解的合法性、与目标的接近程度以及从多个可能解中选取最佳的方案。通常采用多重准则(criteria)法来解决“评价”,从而可以看出该模型的三个阶段具有顺序的、循环的特点,每一次循环不是简单的重复,而是比前一次来得更为详细,因此该模型可以表达成为图 5.2 的递归形式。

图 5.1 设计活动的三个阶段

图 5.2 三个阶段的递归形式

由此可见,在综合开发前没有必要将设计问题完全分析清楚,这一点符合人们的设计习惯和设计的实际情况。

2. 生成-测试模型

生成-测试模型(GT)是由 Popper 于 1972 年提出的,其主要观点是将设计活动视为在一个状态空间中的问题求解搜索的过程。首先是生成一种假设,然后用已有的现象或数据去测试,如发现有不能满足假设的现象,则再次生成一个假设,如此重复,直到找到能符合所有现象的假设作为设计的解。

我们知道,设计类问题大多是一个病态结构(ill-structure)。为了使用该模型,Simon 于 1973 年提出,设计师通过将原始的设计问题降级成有组织的一组子任务,从而使病态结构转化为良性结构(well-structure)。一个设计师,可以随时从其长期记忆中回想起某种约束或某个子目标,但所有这些因素却无法包含在问题描述之中,所以问题求解中任务的形成是动态的。对问题的描述应不断地进行修改,以解释其真实情况,因此问题求解器需要面对的是一个良性结构。

可见该模型简单明了,因此许多设计系统中都或多或少地使用这种思想。但对于复杂的设计问题,由于难以实现问题结构从病态到良性的转移,因而就很难整个地使用此模型,也就只有将它用于局部的范围之内。

3. 约束满足模型

约束满足模型(CS)的出发点是把设计形式化,以逻辑表达设计要求(即对设计问题的描述),通过逻辑推理的办法得到最终的设计结果。它把设计的最终要求概括为一组特性以及相应的约束条件,并以此作为问题求解的最终状态。设计任务从初始问题状态开始,每一中间状态中都包含这些特征,其推理过程是不断满足状态中特性的各个约束条件。基于约束满足模型的设计是当前最为流行的,第 4 章中已经介绍了许多的参数化设计的概念和方法,这里不多重复了。

4. 基于知识的设计模型

CAD 的知识工程方法是一种基于知识的设计模型。它把设计师的知识提炼出来构成知识库,并通过对知识的运用来进行设计,通过知识的学习来改善知识库的内容,提高系统的设计能力,所以称之为 CAD 的知识工程方法。其中最为成功的便是专家系统设计模型,它的设计问题知识库常被分成两类:设计过程的知识,即关于如何进行设计的知识,其中包括设计一般原理、设计的常识等;设计对象的知识,即设计对象的部件、结构、材料、用途、设计规范、典型产品、结构原型和部件类型等。

基于知识的设计模型主要有两种策略:第一种策略是让计算机复制人类的设计行为,仅仅是让计算机进行领域的某项设计,但由于对具有智能行为的设计的研究并未彻底清楚,使得当今的知识表达、自动推理以及问题求解技术只能迎合极小部分的人类设计行为,专家系统便是朝着这种设计专门化的方向所作出的努力。第二种策略是借助于智能工具,为设计人员提供智能支撑。这一方法不仅缓解了设计研究中可用手段不足的局限

性,而且使得我们能在更大的规模及更高的复杂性层次上去研究设计中的智能活动。爱丁堡大学的 EDS(Edinburgh Designer System)代表了这一领域。他们认为目前要提出一种完善的设计理论为时尚早,因而提出了一个基于探索的设计模型,用以作为对设计的智能支持,该设计模型如图 5.3。图中,DKB 表示领域知识库, K_{dm} 表示领域知识, K_{dn} 表示设计知识, R 表示对初始设计要求的描述, E_d 表示设计探索过程, H_d 表示设计探索历史, R_f 表示最终设计结果, D_s 表示最终设计说明,而 DDD 表示设计描述文档。该模型中,知识库是动态的,设计的探索过程以及设计的历史状态将不断地引起领域知识库的增值。同时,新的知识库也影响设计的过程,整个设计是在不断探索中完成的。

图 5.3 EDS 设计模型

5. 设计思维模型

上述各类设计模型存在着一个共同的缺陷:它们并未从人脑认知思维过程的深层去研究设计问题(或仅仅是简单的认知模型而已)。因此,尽管人们绞尽脑汁地提取设计专家的知识,但由于这些知识的运用与人之真正认知过程相差甚远,而使得计算机的设计模拟并未真正地体现出人类的智能。从而必须从研究认知、思维出发,然后建立反映设计思维本质的设计模型系统。

浙江大学人工智能研究所研究形象思维为突破口,进行了设计思维模型的研究。认为设计思维过程远远不止推理、比较和搜索这类抽象的思维操作,更重要的是诸如联想、变形、综合等形象思维类的操作,通过时空、情感相关、概念类似、感觉特征类似等导航机制来完成从记忆网络中的一个结点到另一个结点的发展的操作。

认知对设计思维模型的作用主要以两种方法进行:第一种是用理论研究的方法,分析综合认知科学,特别是形象思维的以形象为核心的形象信息模型;第二种方法是用实验心理学的方法,研究设计思维过程的模型。该方法指出了设计的多模型特性,并分析出形状方案设计思维的四种模型,它们是:对象先例型、约束联想型、分解综合型和抽象逆反型。

5.1.2 基于 CASE 的设计模型

在智能 CAD 系统的设计模型中,根据人类解决问题时往往都借助于先前的某种例子,也就是问题求解时,借助于以前求解类似问题的经验或方法来进行推理,这样做可以不必从头开始。这就是我们熟知的 CASE 推理,它是近来发展十分迅速的一种设计模型。

所谓 CASE 是计算机辅助软件工程(computer-aided software engineering)的缩写。它是建立在基于分析与设计的方法学上,采用良好定义的技术,包括结构化、代码生成等

技术。它不仅支持分析与设计阶段,也支持开发阶段,所以它是一个分析与设计工具,它不是只遵循一个专用的程序设计方法,而是遵循一个良好的定义的,并通常是认可的过程。

我们知道 CASE 是根据当前的问题从 CASE 库中检索出相应的 CASE;调整该 CASE 中的求解方案,使之适合于求解当前问题;然后求解当前问题,并使之形成新的 CASE;最后根据一定的策略将新 CASE 加入到 CASE 库中。在设计中,成功的经验则被用来指导当前的设计,以前的失败经验被用来避免犯类似错误。可以看出,设计活动中的推理过程与 CASE 推理非常相似。本节以 CASE 推理作为出发点,选择了一种非常典型的设计活动造型设计为研究领域,分别从 CASE 的表示、CASE 的组织与检索、CASE 的调整与学习等几个方面来探讨智能 CAD 系统中的 CASE 推理技术。

1. 基于 CASE 推理的信息表示

在人工智能的基于 CASE 推理问题的求解系统中,典型的 CASE 一般包含问题的说明信息(问题的初始条件)、问题求解的目标、达到该目标的解决方案等三部分信息。

由于设计问题一般比较复杂,有关这三部分的信息很难用单个的 CASE 来表示。况且,在实际的设计活动中,有相当一部分的设计问题都是定义不良的,很多表面上一致的设计要求实际上隐含着冲突,给设计 CASE 的表示带来了困难。所以一般来说,设计过程可分为以下两个阶段:概念设计阶段,即进行方案的生成和方案落实阶段;方案落实阶段,即根据设计方案进行设计结果的生成操作,形成最后的设计结果。

对模型设计 CASE 的表示分成两部分:一部分为概念设计 CASE,用来支持设计活动中的方案设计;另一部分为形体设计 CASE,用来支持实际形体的生成操作。概念设计 CASE 中包含有设计问题的初始约束条件、设计所要达到的目标(即设计要求)以及满足设计要求的一个可行的设计方案;形体设计 CASE 中包含有设计方案、形体间的约束要求、表现手法、模型操作序列及一个实际的模型设计结果的描述等。之所以给出设计结果的描述是因为在设计活动中,设计要求与最后设计结果往往有一段“距离”。下面选用模型设计中较为常见的室内设计为例来作进一步的说明。例如对于一个单间居室的室内设计,其概念设计 CASE 和模型设计 CASE 可分别表示如下:

(1) 概念设计 CASE

房间说明: 单间、面积、长、宽、高等;

设计说明: 居住者说明,即性别、年龄、职业、爱好等;

功能说明: 提供学习、会客、体息等功能子空间;

设计要求: 指标说明要求经济性、美观性、实用性、私密性等;可由写字台、转椅、书柜等构成一个学习子空间;由床、衣架、衣橱、床头柜等构成一个休息子空间;

设计方案: 由沙发、茶几、折叠桌等形成一个会客子空间,子空间之间的分隔可采用软分隔或硬分隔。

(2) 形体设计 CASE 设计方案可以同上面一样。

实体间约束要求: 子空间之间的几何约束及其各组成部分的约束,可以采用地台来增加视觉层次;

表现手法:放置植物,以显示生命活力;用房间的摆设来显示主人的性格、爱好等;构造书橱的操作序列、构造沙发的操作序列、构造写字台的操作序列等,模型操作序列、书橱的放置操作序列、沙发的放置操作序列、写字台的放置操作序列.....。

其设计结果是使该 3D 房间模型的数据结构由 B-rep 或 CSG 及一些说明信息来表示。

通过上例说明,大致了解了设计 CASE 中应包含的信息,其概念设计 CASE 以“影响图”的方式来表示,将其中设计说明、设计要求及设计方案的每一组成部分看成一个节点,根据它们之间的信赖程度用加权有向弧将它们连接起来,形成一个影响图。对形体设计 CASE 则采用分布式的 CASE 表示,将其分割为很多小的片断 CASE,这些片断 CASE 围绕特定的设计目标而组织,通过片断 CASE 链接来保持整个形体的结构信息。

2. 基于 CASE 设计的组织与检索

CASE 的组织与检索是进行 CASE 推理系统设计的一个重要环节,其策略的好坏对 CASE 推理的效率有很大的影响。在 AI 中,CASE 的组织与检索策略主要有以下几种:最近相邻策略(nearest neighbor),它根据 CASE 中各组成部分的权值,求得其权和,然后根据其权和的远近来组织相应的 CASE,检索过程与此类似;归纳推理策略,即不断地从 CASE 的各组成部分中抽取出最能将该 CASE 与其它 CASE 区别开来的成分,并根据这些成分将 CASE 组织成一个类似于判别网络的层次结构,检索时采用判别树搜索策略;知识导引策略(knowledge guided),它根据目前已知的知识来决定 CASE 中哪些特征(或信息)在进行 CASE 的检索时是重要的,并根据这些特征来组织和检索,这使得 CASE 的组织与检索具有一定的动态性。

从知识使用角度出发,虽然有些系统对 CASE 的检索与存储技术作了一些优化,如抽象索引、部分匹配、泛化匹配等,但总的策略并未超出上述三类。那么,设计 CASE 的组织与检索策略应该是:在实际的设计活动中,设计师往往随着其对“设计世界”的看法的变化而动态地组织其设计知识。对模型设计 CASE 的组织而言,这种动态性主要表现在设计方案与实际形体的对应上受不同的审美观、表现手法及设计习惯的影响。这就要求模型设计 CASE 的组织也必须在这方面表现出一定的动态性。而目前 AI 中的 CASE 组织策略不能简单地应用到设计 CASE 的组织上,为此必须采用以下策略:

第一种策略是对概念设计 CASE,由特定领域,如模型设计的概念语义模型来组织它们。首先通过概念模型对它们进行一个大致的分类,然后对还不能区分开来的概念设计 CASE 采用最近相邻策略来组织,从而使得概念设计 CASE 中隐含有语义成分。这对以后 CASE 的检索与调整十分有利。

第二种策略是对形体设计 CASE,采用归纳推理和知识导引相结合的策略来组织。首先依据归纳推理对形体设计 CASE 作适当的区分,然后由一个有关方法的选择、评价等方面知识的具有学习功能的导引器来组织。这样通过导引器中知识的演变来实现设计 CASE 的重组,在一定程度上体现了动态组织设计 CASE 的思想。

从 CASE 的组织策略可知,它已在相当程度上决定了 CASE 的检索策略。CASE 检索操作的目标是快速地从 CASE 库中找到一个跟当前的问题求解最相关的一个(或多

个)CASE。但对设计 CASE 的检索来说,仅由 CASE 的组织策略来进行相应的检索操作是不够的,这是因为设计活动中的知识使用具有“演变性”。这样的“演变性”决定了在进行设计 CASE 的选择时,不仅与相似性确认程度和子行为匹配程度有关,而且还与该设计 CASE 的被使用的“历史”有关。针对设计活动的这一特性,我们在概念设计 CASE 中增加了一个组成部分,用来记录 CASE 被选中的成功或失败的历史,以利于最佳 CASE 的确认;而对形体 CASE,因其采用分布式的表示,故对其各片断 CASE 保留其被调整成功成功率。这为以后的 CASE 调整操作提供了很好的支持。

3. 基于 CASE 设计的调查与学习

为了在检索操作中获得最佳 CASE 中的求解方案,必须进行适当的调整操作,使之适合于求解当前问题。由于整操作与特定的领域相关,因而也很难定义一个通用的方式来执行;同时,也为了解决设计要求中的定义不良问题,应对设计问题中隐含的约束进行一定的整合与协调。例如,因为已把设计 CASE 的表示分成概念设计 CASE 和形体设计 CASE 两部分,对形体设计 CASE,其调整操作如下:首先隔离出可用的片断 CASE,然后对不合要求的片断 CASE 采取功能相同替换、功能分割等操作来进行调整;再根据片断 CASE 间的链接信息进行重构,并形成一个新的 CASE 加入到 CASE 库中。

随着新的 CASE 不断地加入到 CASE 库中,将会使 CASE 库显得十分庞大,其推理效率必然受到影响。为将 CASE 库控制在一定规模内,必须对加入到 CASE 库中的 CASE 进行学习,这样的“学习”实际上就是把多个 CASE“合并”为一个。目前 AI 中 CASE 学习仅是一个泛化过程,即只进行相似性学习。但是,对设计 CASE 学习来说,仅从相似性角度来进行学习是不够的。我们认为,它还应能够根据设计 CASE 的被使用情况进行学习,这才能体现出设计知识的演变性。

4. 推理过程

根据上述对 CASE 的表示、组织、检索、调整、学习等讨论,可以得出一个完整的基于 CASE 的设计推理过程如图 5.4 所示。

其中初始条件和设计要求是通过接口转换成一个概念设计 CASE,然后由概念设计

图 5.4 CASE 设计的推理过程

CASE 检索器从 CASE 库中搜寻出最相关的 CASE,由调整器对设计要求中的隐含约束进行整合,并执行相应的方案调整操作,形成一个可行的设计方案。根据设计方案从形体设计 CASE 库中找出相应的设计 CASE,分隔出其中有用的片断 CASE,并对不合要求的片断 CASE 执行调整操作。调整成功后,则执行形体设计 CASE 中的实体模型操作序列,形成一个实际的形体输出给用户,同时将形成的新的 CASE 通过学习器加到 CASE 库中。为便于理解,我们把概念设计 CASE 的推理和形体设计 CASE 分开讨论。实际上,在概念设计 CASE 时,可以通过泛化操作来访问方案落实 CASE 中的知识。相应地,在进行方案落实 CASE 的推理时,也可以通过泛化操作来访问概念设计 CASE 中的知识,这样的设计推理过程是“非线性”的。

5 2 智能 CAD 系统的设计方法

智能 CAD 系统的设计方法很多,如 CAD 模型自动识别的神经网络方法、基于几何约束的表示满足方法、不正确性推理、不精确推理、几何推理以及现有的智能工具如专家系统工具等,以下仅介绍面向对象的求解方法、设计过程模型的特征设计方法以及广义类比方法等。

5 2 1 面向对象的求解方法

求解设计模型的一种较好的方法是面向对象的求解方法(ODPSM),因为设计过程是体现人类智能的、具有创造性的行为,它所涉及的领域很广,各个领域的设计方法又有各自不同的特点,这使得建立计算机辅助设计系统的困难大大增加。以下介绍的 ODPSM 模型及其推理机制正是对计算机辅助设计系统的知识表达和推理方法的一次探索。它的类表示能完成一定功能的问题求解器,同时也是设计对象的抽象描述。类间关系由 Used-By 域和复杂属性来表达,且在求解过程中才建立实际联系。支持推理的知识分成两部分描述,即求解方法和约束满足。

现有的 CAD 系统多侧重于图形显示功能,但对真正的设计过程缺乏支持。具体体现在智能性不够、用户界面不友好、集成程度不够、信息交换困难等。对上述不足,STEP 标准的实施为不同模块间的信息交换提供了可能。可见,最关键的问题仍是 CAD 系统的智能性。在处理 CAD 系统的智能问题时,有两种不同的方法可采用,即基于专题的系统和基于模型的系统,而 ODPSM 就是这样一种设计系统模型。它提供了多种知识表达方法,并以面向对象的方法根据问题的需要进行结构化地组织,并在此基础上建立适用于设计问题求解的推理机制。

1. 面向对象求解方法

我们知道,一个设计模型描述了设计对象的各种信息和设计知识,如属性、结构、相互关系等,更重要的是如何表达体现人类设计专家智能行为的知识。这类知识有它不同于一般专家系统知识的特点,它们不仅仅是设计专家经验的结晶,更多的是在设计过程中要遵循的一些设计规范和要求,而且在某些领域中,还有在相应设计理论指导下的一整套切

实可行的设计方法,人类专家的设计行为也是在这些理论和方法框架之内的。因此,一个用于计算机辅助设计的求解模型应充分考虑到这一点。现有的设计理论和设计过程模型,以及支持设计过程的 CAD 系统原型,有一种认识上的共同点,即认为设计系统应该以结构化的方式来组织,而且在多数情况下为层结构,这也体现了设计过程由顶向下的不断精化过程,一个大问题被逐级分解为子问题,子问题的求解是相对独立和互相制约的。现有的设计系统原型多数是基于框架结构或类似于框架结构的,这种结构对于参数设计是比较合适的。通过约束满足来达到对象间相互制约的要求。

(1) 类的定义 在面向对象的设计问题求解模型(ODPSM)中,设计问题被认为是可分解的,具有一定的层次结构特点。每一个担负一定设计求解任务的单元可以定义为一个类,这是相对独立的问题求解器,它所能完成的功能可以是从一个设计参数的确定直到整个设计结果的综合中的每一个可能的步骤。ODPSM 类的每一个域都赋予一定的语义,以满足设计问题的求解需求。

(2) 设计知识的表达 ODPSM 中的设计知识是由求解方法和约束要求两部分组成。其中求解方法是在 ODPSM 的 Action 域中定义了当前求解器在求解过程中的行为序列,它的主要任务是设法使求解器中的属性值得到确定。在每一个问题求解器中,按具体情况明确地定义了该求解器在求解时的动作序列。它的每一个动作可以为以下几种操作:首先是 Action 序列本身具有一定的简单计算能力,可对一些简单的操作做出决定;另外,它提供了 If-Then-Else 的控制结构,以满足求解过程中必要的分枝要求。其次是规则推理的启动,产生式规则用于整个设计问题的描述虽不合适,但产生式规则作为一种有效的知识表达方式,有它独特的优点,如在求解一个有多种可能的小问题时,它能够很好地胜任。因此,ODPSM 也把产生式规则作为一种知识表达方法,还有外部过程的调用。ODPSM 采用外部函数调用的形式将这些常规技术(如分析方法等)结合到系统中来,Method 定义了外部函数的接口原型。最后是复合对象的求值,这种操作实现了设计问题的分解和子问题设计方案的选择。

约束要求是 ODPSM 中的一种重要知识表达方式,它的主要作用是描述不同对象间的相互协调关系和所属求解器在求解过程中要遵循的一些一致性约束。ODPSM 的约束集由一系列约束项组成,每一项是一个独立的功能单元。

2. 推理机制

一个智能系统的智能性体现在它的推理行为上,同样地,一个计算机辅助设计系统所能表现出的智能行为,关键在于不同领域中采用合适的推理。

设计问题是一个很复杂的问题,主要体现在设计对象各部分之间的关联很多,有不少是相互制约的,而且不同领域的表现形式也有较大的区别。因此,几乎不可能要求计算机用一种通用的算法解决设计问题。对于设计问题,一个较为有效的方法是从问题的分层结构出发,对每一个子问题提出当前的可行的解决方案,当以后某个时刻发现这个方案和其它部分的设计发生矛盾冲突时,再进行方法的每一步操作,不管是子问题的求解、矛盾的判定,还是发生冲突后的调整方法,都有很强的应用背景,不同领域的方法可能是完全不同的,它体现了不同领域的特殊设计知识和方法。因此,整个系统的行为应该在很大的

程度上信赖于应用领域的知识,由具体知识指导实际的推理流程。ODPSM 推理核心提供了基本框架结构和功能,以支持这些方法的有效使用。

图 5.5 就是求解的流程图。当一个问题求解器被启动时,系统完成一系列初始化,主要是产生一个新的实例工作区,并赋以惟一标识,然后根据 Action 定义的内容进行逐步求解,直到所有设计指标的值都得到确定,或是整个 Action 动作序列中没有新的操作要求。在求解器完成工作之后,将此次产生的实例存放于实例数据库,然后根据情况向调用它的父求解器汇报,并唤醒父求解器。

图 5.5 面向对象设计问题求解流程

根据实例关系的不同和传输操作的复杂性,可分为继承、引用和消息传送等三种方法。其中继承只能用于实例树型结构中的祖先子孙关系中,子孙可以继承祖先的全部设计结果信息,引用用于求解器层次结构中的父子关系中;消息传送是当两个需要传送信息的求解器间不存在上述关系时,而只能采用消息传送某种信息的请求,后者接到请求后,将相应的信息拷贝一份送回。

约束满足检查是自动进行的,当某次操作引起部分属性值发生变化时,就进行相关约束检查。为了能高效地实现这一点,必须对每个求解器的约束集建立相应的网结构。网中的节点有两类,一种是属性节点,它描述了当前求解器的各种属性和其它求解器中的非私有属性,这是引起约束检查的源点;另一种节点是函数值节点,它描述了一个求值函数,可能是系统包含的简单函数,也可能是用户定义的特殊函数。约束关系由带标记的弧来表示,除了系统定义的约束类型以外,还有可能是由用户自定义的以谓词形式出现的约束关系。标记上同时包含有约束不满足时进行调整的方法。约束不满足时的调整操作根据其作用范围可分成两类:一种是对当前求解器内部简单属性的调整,这可以在现场直接进行;另一种是需要其它求解器来完成的调整,在处理这类操作时,系统将当前求解器挂起,

同时向相应的求解器发出请求,当前求解器在相应的请求完成以后被唤醒,然后根据递交的结果进行进一步的操作。

5 2 2 智能 CAD 的广义类比推理方法

类比推理作为一种思维方式和知识获取工具,正愈来愈受到人们的重视。在日常生活中,人们常常要用具体行动的例子来代替枯燥乏味的说教,才能取得良好的效果,尤其当对象缺乏理性和显式的结构性时,更是如此。研究表明,类比学习是人类认识初期过程中的主要方式。实际上,类比是形成假设的主要来源之一,许多人类的高级成就萌芽于类比。虽然类比是介于理性与非理性之间的一种思维方式,并不能保证其结论永真,但作为一种贴近人类思维的模拟方式,开拓思路,发掘新事物,具有不可替代的优越性,其潜在的应用价值亦不可估量。法律上的类比断案、机器故障诊断、自动程序设计都是类比推理的狭义应用;从广义上讲,从纵横表得出一枚别针的几十种用途,由概率前提 A 和 B 得出概率结论 C , 实际上都是类比性的推理方式。

1. 类比推理概念

客观世界是一个存在着普遍联系的统一有机体,各处貌似迥异的事物之间都或多或少地存在着相互联系,这种联系是内在的必然的。应用类比推理,利用这种普遍联系,就可以从现象(前因)找出服务于特定目的的结论(后果)。这就是我们对这个问题的基本着眼点。

我们可以认为:若事物 A 和 B 在某方面是相似的,则称 A 和 B 是可类比的。这个定义符合人的直觉,一般当两个事物存在着共性时即认为它们是同类的,所以又可以认为:若 A 和 B 是可类比的,则在它们得出可类比性的属性以外的其它对应方面,也存在着可类比性,这个过程称为类比推理。在这两个定义的约束下,类比推理存在于已知的界限以内的事物之间,推理的过程只是从事物的一个方面扩展到其余方面,它并不带来新鲜内容,至多只是提示被忽略的方面;对陈旧的知识进行有限的改良,其最后结果亦不外乎是固定模式下遗传的积累。为了进行更真实的推理,我们拓广类比的概念,以赋予类比推理继承性以至创造性。

我们还可以认为若事物 A 和 B 在任何有意义的方面是相似的,则称 A 和 B 是广义可类比的。而且若 A 和 B 是广义可类比的,则由 A 和 B 的已知可类比推知其余方面的可类比,并在此可类比意义约束下产生与 A 和 B 皆为广义可类比的异于 A 和 B 的子代 C 的过程,称为广义类比推理。

因此,从上述两个定义,得到了类比推理的全新意义,它不再局限于已知事物而产生全新类比物 C 。 C 受 A, B 的类比约束而又不同于 A 和 B ,所以这种广义类比是遗传和变异的结果。由于有遗传,所以 C 与 A, B 可类比;由于有变异,所以 C 具有创造性。与原先的定义相比,这种广义类比是进化性的。

2. 推理过程

在进行广义类比时,一般要确定类比对象、类比空间、类比范数。

类比对象是推理的物质基础,在确定类比对象时,首先要确定推理方向,若是正向推

理,则对象是 A 和 B , 结论是 C ; 否则就应出发于 C 而止于 A 和 B , 进行逆向推理。正向推理对应于看到两样事物想象结合两者会产生什么; 逆向推理则对应于看到一样事物去想象它与哪两样东西有共性。

类比空间由对象的特定属性组成, 因而它是对象所有有意义属性的幂集, 这些属性包容各个方面如颜色、空间关系, 也可是刺激度、历史传统等等, 其选择依赖于应用的衡量准则。这些属性, 大致可分为语义属性和结构属性两类。

类比范数是在已定的类比空间中, 进行怎样程度的类比, 其自由度受限于类比范数的确定, 这个范数的选定同样依赖于特定问题, 比如它可以是欧氏距离, 也可是矢量积叉乘等等。经过类比对象、空间和范数的选择, 类比即可在一定程度范围内进行, 它既避免了演绎推理的封闭和僵硬, 又在一定程度上接纳了模糊不定的概念准则。

实际上, 类比的双亲可以是一组 A_1, A_2, \dots, A_N , 此时正向推理即变为创造揉合了 $A_1 \sim A_N$ 风格的后代 C 的过程; 而逆向推理转化为搜索类比空间中在范数约束下的所有有关对象。

3. 类比推理的概念模型

根据前面的讨论, 类比推理的模型应是一个三元组:

$$M = (\text{Object}, \text{Attribute-space}, \text{Form})$$

其中, Object 是类比对象, 隐含了推理的方向, Attribute-space 是类比属性空间, Form 是类比范数。

图 5.6 是这个模型的推理过程的一个简单示意图, 在图 5.6(a) 正向推理时, 各个 C_i 均来自 A 和 B , 但 C_i 又各不相同, 它们可能位于同一范数约束下轨迹的不同位置, 也可能位于不同范数所确定的不同轨迹上。介于 A 和 B 之间的 C_i 是由 A 和 B 插值所得, 在 A 和 B 之外的则是以 A 根据 B 或以 B 根据 A 外推所得的, 显然这两类结果的几何意义是不同的。内插是一种揉合过程, 而外推则是一种创造过程, 同样在如图 5.6(b) 所示的逆向推理过程中, 由 C 可以在不同的类比空间和类比范数下得到一族 (A_i, B_i) , 它们各不相同, 又都在特定方面与 C 可类比, 这实际上就是一个充分发挥联想的过程。

图 5.6 类比模型的推理过程

正是由于内插和外推, 使得这种广义类比具备了巨大的生产力和创造力, 能产生很多新型的结论对象, 这对于开发新产品, 进行发散式思维都有重大意义。但值得注意的是, 过分外推即是“胡思乱想”, 会陷入混沌状态。

在计算机图形学中,将上述模型广泛应用于漫画创作和动画制作。如前所述,事物之间存在着普遍联系,所以可将一个小兔与一个老鼠作类比,图 5.7 所示为内插得到的小兔向老鼠的“渐变”动画过程。

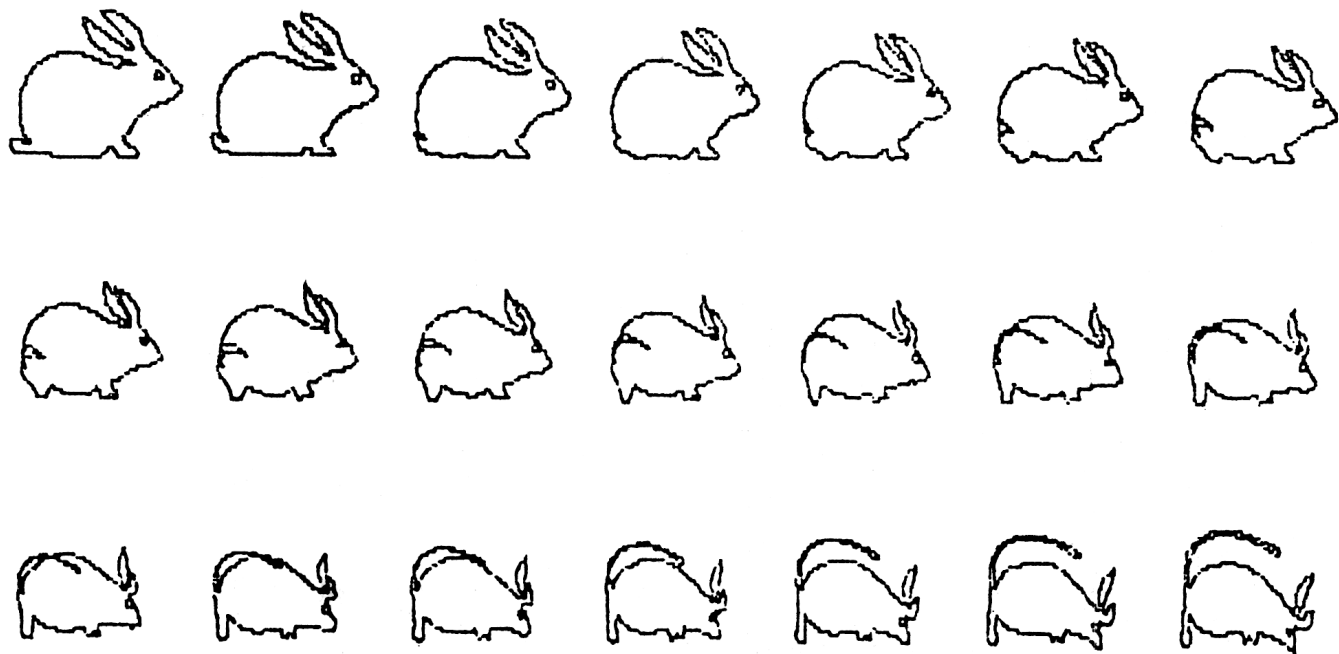


图 5.7 类比应用实例

在广告设计中,这个类比模型可用于嫁接两幅图画,如合成一帧纸杯可乐照片和一帧一杯冰镇玻璃杯可乐的照片,结果是一帧下半截为玻璃杯的可乐形象,这种效果是在摄影或暗房中难以实现的。

总的来说,在将上述广义类比模型应用于新产品开发过程中时,只要类比对象充分多,属性选择恰当,就可以产生千姿百态各种新品,即使一枚别针也可发生几千万、几亿种变形应用。

5 2 3 设计过程模型的特征设计方法

产品造型的设计过程理论主要是研究设计对象的建模、变换及相关技术,它是 ICAD 的理论基础。对 CAD 系统智能性判断的重要标准也就是系统对设计过程的认识和建模。从应用角度来看,有自上而下设计、并行设计等过程模型。从第 4 章的介绍可知,这些模型中的特征的概念与建模是信息组成的基本单元,也就是把设计描述成“单元”及其联系,把设计过程视为模型的渐进过程,而这时的信息“单元”被称为广义上的特征。由此可以这样认为,设计是一个基于特征变换的过程,是对基于约束的设计过程模型之补充。

1. 一般设计过程

一般设计过程分为需求分析、概念设计、初步设计和详细设计等阶段。以图 5.8 的零件设计为例的需求分析是指零件在其工作环境中有两个主要功能,支撑一个轴,并定位于指定位置。概念设计是指功能需求被映射为设计对象的原型结构,如图 5.8(a)所示。原型的选择或产生是一个模式匹配、类比与修改的过程,既可以是创造性设计,又可以是修改设计,既可以视零件为基本单元,又可以将特征视为功能和结构要素。初步设计是指对

原型的“加工”可能有两个步骤:先进行结构的选择和变换,消除原型的不确定性和不完备性;再对初步模型进行工艺、性能等诸多评估,并进行基于知识的修改设计,如图 5.8(b)应增加一个加强筋,定位通孔要改为沉头孔。详细设计是指对图 5.8(c)进行工艺化设计得到图 5.8(d)。

图 5.8 零件的一般设计过程

2. 设计过程模型

设计理论要确立其科学地位,就必须对设计过程进行建模。产品模型要最终超越几何模型而成为 CIMS 的核心技术,也必须对设计过程作深入研究。因此这里将把产品模型建模过程表述为基于特征变换的过程。

在第 4 章已简单介绍了特征一般定义与基本分类,设计过程中的特征定义与分类,应根据特征的概念,根据应用对象和应用领域不同而有所不同。在设计过程中,分析已有的特征定义,认为形状特征是产品全体或某一局部相关的几何知识、CIM 领域和 CIM 环节知识的三元组。这些知识具有特定组合、生成与变换规则,能够用于设计、识别和分类。就整个 CIM 来看,形状特征具有动态特性和相对稳定性,比如“孔”可以有很多变体,但其原型不变。形状特征还有其自身生命周期,未必存在于产品的全生命期中,可以根据形状特征在设计过程中的表现进行分类。

根据特征在设计、工艺、分析等环节的表现,特征又可分为设计特征、工艺特征、制造特征、分析特征,诸如此类,这就要研究特征映射。根据特征的组成和功能特性,及其在设计过程中发生作用的方式,又有简单特征、复合特征和特征组之分(图 5.9)。产品的某些功能要由一个甚至多个零件上若干个分离的特征来共同完成,这些特征在设计过程中要被视为一个整体,有自身变换方法。如图 5.9(d)是“轴段、键槽和轴肩”这一特征组的变换,轴、键和齿轮由齿轮轴来代替,这样便减少了零件数,简化了装配和制造,因此根据系

统化设计的需求,每一类功能都可以由特征族来实现。对于变量设计常常需要采用相似特征之间的替换,这是特征变换的另一种形式。

特征分类的另一依据就是特征的完备性及其表现,大致可分为原型特征、设计特征、技术特征、工艺特征等类。

产品原型中的特征统称原型特征,这是对形状特征的高度抽象。例如,孔类特征表现为圆通孔甚至中心线。图 5.8(a)特征 F_1 , F_2 是功能结构映射的结果,特征 F_3 则处于从属地位,因为其几何属性取决于 F_1 和 F_2 。

初步设计中的形状特征是对原型特征的具体化,称设计特征。如图 5.8(b)中的中心线、轮廓线这样的抽象被具体实现,以保证初步设计是无歧义的。当然,初步设计的性质、尺寸方案是可变的。

设计特征还要进一步变换,例如,设计特征的细化、组合和变换,或者某些特征具有自身的设计规则,由此引入的特征称技术特征。如图 5.9(d)的特征变换,图 5.8(c)的加强筋都是技术特征的表现。

设计的最后阶段要用倒角、圆角之类修饰特征,用于对已有特征进行修改,故特征的最终表现称为工艺特征。

图 5.9 特征的组合

(1) 设计过程中的特征变换 设计可视为是特征变换的过程,上述的每一种特征分类方法都对应于一类特征变换。这里把特征变换概括为特征的产生、继承、变异、抽象、派生、映射和消亡。特征产生是指因功能、工艺等需要而创建特征;特征继承是指产品模型变换中继承前趋模型的特征,诸如类型、尺寸继承;特征变异体现了特征的渐进特征,例如,特征的分解、组合和修改;特征抽象,例如把孔抽象成中心线;特征派生是特征抽象的逆过程,在相似设计中极为重要;特征映射体现了从功能、形状、制造等多视角来认识产品模型;特征消亡是指特征相交、合并或制造装配的分析导致特征自动消失或被有意删除。

(2) 设计过程模型 图 5.10 设计过程模型,是依据特征分类将设计过程分为需求分析、原型设计、原型细化、技术设计和详细设计等环节,基于知识的特征变换贯穿始终。图 5.10 的简化模型略去了回溯与评估。若用 $P(F, C)$ 表示特征 F 在约束集 C 下被组织

成模型 P , 那么原型设计

$$P_0 (F_0 , C_0) = \varphi_0 (k_P , k_F , R)$$

把功能规范 R 映射为原型特征集 F_0 和约束集 C_0 , φ_0 依赖于原型知识 k_P 和特征知识 k_F 。原型设计实际上就是概念设计。

而原型细化

$$P_1 (F_1 , C_1) = \varphi_1 (k , k_F , P_0)$$

继承和例化原型 P_0 , 并根据设计知识进行 F_0 的派生、复合、分组等。

技术设计

$$P_2 (F_2 , C_2) = \varphi_2 (k , k_F , P_1)$$

根据设计知识和“为制造而设计(DFM)”、“为装配而设计(DFA)”等哲理, 进行特征和特征组的创建与变换, 得到技术模型 P_2 。

详细设计

$$P_3 (F_3 , C_3) = \varphi_3 (k , k_F , P_2)$$

对技术模型作工艺化设计, F_3 就是工艺特征。

由此可见, 可由几何求解, 最终得到模型的边界表示。上述模型具有如下特点: 特征变换是设计的基本活动, 其设计具有特征驱动的性质。特征可能提出相应的设计要求, 从而影响后续过程; 制造、装配等的需求都是特征变换的内在动因, 所以设计具有“Design for X”特性(“X”即制造、装配之类)。

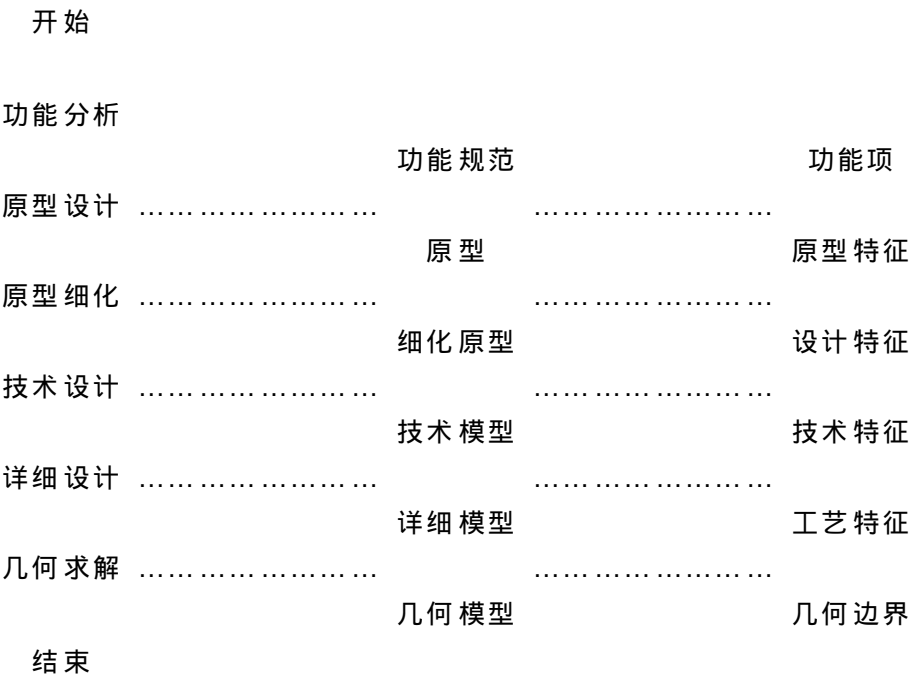


图 5.10 设计过程模型

3. 应用

通过记录各阶段的产品模型及其特征变换过程, 可以解决许多应用问题。总的思路是对不同应用选择合适的阶段模型。类比设计要用到产品原型, 而力学分析、装配描述等最好不要基于详细模型, CAPP 和尺寸标注则应利用产品信息的层次性。下面举三个例子:

(1) 类比设计 目前 CAD 系统仅能做到参数设计。在我们的设计过程模型下,通过特征抽象就能得到产品原型,再以人机交互方式或者自动方式进行特征变换,就可以派生出各种各样的相似模型而不仅仅是拓扑同构的边界表示。

(2) 装配描述 产品模型具有很大的可变性,而原型则相对稳定。图 5 .11 所示为轴、齿轮和键的装配关系,联接特征指向边界表示。若修改轴的设计,如增加过渡轴段或中心孔,有理由认为,修改前后的装配描述是相似的,但现在的描述方案却不如此。如果把装配关系抽象为零件原型之间的联接,装配表示就能表示一族产品的装配方式,所有装配关系随特征变换而变换。

图 5 .11 轴、齿轮和键的装配描述

(3) 尺寸标注 一般主体尺寸取决于产品原型,细节尺寸来源于特征变换。而采用基于参数化设计和基于特征的设计过程就有可能实现自动标注尺寸。

以上讨论只给出了设计过程的概念模型,其实还有许多问题需要讨论。就 CAD 系统实现来看,设计过程模型还要进一步形式化,包括特征、产品模型的表示,还有特征变换,实际上特征变换隐含了特征之间约束关系的变化。

5 3 智能化空间布局设计

空间布局设计研究是智能 CAD 的一个重要领域,早在 60 年代,研究者就开始了计算机辅助建筑平面布局设计的研究,然而目前尚未取得十分令人满意的结果。Montreuil 指出,计算机辅助布局设计的方向在于所谓“智能布局设计环境”,要形成这种环境需要在领域知识建模、启发式搜索技术、优化技术、仿真建模与内部环境模型交互等多个方面进行艰苦的研究。空间布局设计是基于计算机科学的、人工智能及思维科学的,其研究除了其自身的应用意义之外,也必将促进这些研究领域的发展。本节主要介绍智能化空间布局设计的理论与算法。

5 3 1 智能化空间布局的分类、建模、求解简介

1. 空间布局设计在 CAD 中的重要性

空间布局设计普遍存在于工业设计、生产与日常生活中。在机械结构优化、钢板等原材料下料、集成电路设计、机电设备内部结构安排、厂房设备布置、集装箱内货物安置、产品包装等领域布局设计成为其核心问题。布局设计在这些应用领域占有举足轻重的位置。布局属于非枝节性(non-trivial)问题,布局的好坏直接影响产品的合理性、经济性和安全性。举例来说,在厂房设备布局问题中,设备的布局极大地影响着厂家的长期利益。

研究智能化空间布局设计是为了在布局设计的部分或全部过程中支持设计师的思维活动,力求达到自动或智能交互设计,以减轻设计师的劳动强度,提高布局设计的效率。绘图是传统 CAD 系统所扮演的基本角色,不少研究者指出,绘图仅仅是各种活动行为间的一种通讯手段,并非是设计过程的一个必要步骤,因此,以改变 CAD 系统的角色,使其能够代替设计师的部分思维活动为目的的智能 CAD 系统应运而生。在空间布局领域引入智能 CAD 方法,研究布局设计自动化或者智能交互设计的方法与技术,以此为基础进一步研制相应的智能化布局设计工具,采用计算机支持布局设计的部分或全过程,对于减轻设计师的劳动强度以及提高布局设计的效率具有极其重要的意义。

研究智能化空间布局设计可以克服人工布局的不足,寻求较优设计方案。人工进行布局设计,受人的能力与知识结构的影响极大,对一个布局问题往往只能设计出有限几种方案,通过分析比较,再从中选择一个较好的方案,这不仅仅会遗漏一些更好的方案,而且耗费大量的时间与精力。Roth J.等发现,在建筑平面布局设计中,对于三居室住宅平面布置,采用他们的计算机程序可自动设计出 384 种方案,其中一些很好的方案却超出了以色列国家的设计方案库中的内容。因而采用计算机智能辅助布局设计有可能获得比仅用人工设计更好的设计结果。

同时,研究智能化空间布局设计将进一步促进智能 CAD 的发展。因为具备设计知识是智能 CAD 系统的基本特征,知识是与领域密切相关的,因此对智能 CAD 的研究应当结合具体领域的具体问题来展开。空间布局智能化研究作为智能 CAD 的一个研究领域,一方面受智能 CAD 一般方法的指导,反过来,在空间布局研究中出现的问题又对智能 CAD 的一般理论与方法提出新的要求。因此,两者必然会相互促进。

2. 智能空间布局设计的分类、建模、求解的研究

空间布局设计过程自动化的研究已有 30 多年的研究历史,已取得一定成果。从布局研究的内容来看,主要包括建模、生成及评价三个部分。建模就是研究表达空间及空间关系的计算机模型;生成就是以空间模型为基础,设计自动生成布局方案的计算方法;评价就是对生成的布局方案按着某些设计标准对其优劣作出判断。以下就空间布局分类、空间建模、求解技术三方面的发展作一简要介绍。

(1) 空间布局的分类 根据布局空间的维数,空间布局可分为一维布局、二维布局及三维布局。由于平面布局(即二维布局)更具一般性,且较三维布局简单,故目前已有的

研究成果大多属于二维布局。

根据布局问题的约束情况,可将其分为无(弱)约束布局问题与约束布局问题。所谓无(弱)约束布局问题是指,对布置物体的位置关系无特定要求。这类问题典型的是所谓切割问题(cutting problems),这一问题首先是由 Kantorovich 提出的,后来由 Gilmore 及 Gomory 等若干学者进行了大量的研究。根据被切割对象的维数,可分为一维切割问题(one-dimensional cutting)与二维切割问题(two-dimensional cutting)。一维切割问题研究将一维物体分割为更小的一些片段(每一片有一个给定的长度与值),使得所有片段的值总和最大;二维切割问题可描述为将一个给定的长方形分割成更小的长方形片(每一长方形片有一个给定的尺寸与值),使得分割后的小长方形片的值总和最大。当每一片取与其面积成比例的值时,切割浪费最小问题(waste minimization)可形式化为切割片的值和最大问题。这一问题可被转化为“背包问题(knapsack problem)”。解决此类问题的主要方法有动态规划方法(dynamic programming),树搜索技术(tree-search techniques)及其它优化技术。经过若干学者的深入研究,通过运用模拟退火(simulated annealing)算法、最佳优先搜索等多种搜索策略,以切割问题或装载问题为典型的无(弱)约束布局问题,有了较突出的进展。

约束布局问题是指布局中对布置物的位置关系有特定的要求。建筑环境设计中的布局、厂房设计布局、车辆驾驶室内的布局等多属于约束布局问题。约束布局问题与无(弱)约束布局问题的区别除了有无约束之外,更主要的在于约束布局问题: 问题是非良性定义(ill-defined)的,因此约束常常难以确定,往往要随着设计过程的进行而发生变化;

许多约束知识难以精确表达,难以用计算机模型加以表示; 难以给出方案优劣的精确标准,限制了布局搜索中启发式方法的应用。

约束布局问题的特点决定了它比无(弱)约束布局问题更为困难。研究者自 60 年代就开始了约束布局问题的研究,然而至今尚未取得令人满意的成果。根据活动空间的基本形状可以将布局问题分为长方形(体)布局(rectangulars layout)问题与不规则形体布局(irregular shapes layout)问题。所谓长方形(体)布局是指活动空间及布局区域为长方形(体)或其简单组合。多数建筑环境布局、厂房设备布局、切割下料等问题基本上可以抽象为这种类型,所以空间布局的研究多集中在这一类型。所谓不规则形体布局是指活动空间可为任意给定的多边形形状,其布局区域一般限定为长方形或其简单组合。不规则形体布局是针对切割问题展开的,其应用背景为衣料、皮革、木板、钢板等切割问题。这一问题的主要研究方法是基于形状特征的各种启发式方法。

(2) 空间布局建模 经过国内外学者 30 多年的研究,根据所研究问题的性质,提出了一些空间模型以及相关的设计求解算法。所建立的布局空间模型基本上可分为两类:一类是基于位置坐标的模型,它直接记录布置物的位置坐标;另一类是基于空间拓扑关系的模型,它只记录布置物的相对位置关系。基于位置坐标的模型中较为典型的有离散网格表示和拓扑层次模型。

离散网格表示(discrete grid representation)模型表达布局的方法,它是将布局区域分割成大小相等的网格,各物体分别占据其中一部分网格,对于平面布局,在计算机内部一般可用二维数组表达一个布局状态,即第 i 行第 j 列的网格对应数组的第 i 行第 j 列元

素 $a(i, j)$; 为了区别各个物体以及空间区域, 可给各物体及空间区域分别规定一个数字属性。离散网格模型表达布局具有便于表达复杂形状物体和表示方法简单、直观等优点, 因此许多布局问题的研究采取了基于离散网格模型的途径。其缺点是: 表达布局往往需要使用庞大的数组, 占用很大的计算机内存; 活动空间之间的位置关系没有直接记录下来, 空间关系属性的获取需要不断地对庞大的数组进行扫描、分析和计算, 耗费计算机内存及时间较多。

在研究过程中, 根据不同问题人们提出了离散网格模型的若干变型, 如长方形顶点的离散整数点方法及其改进, 如四叉树、八叉树等等。改进的模型虽然节省了空间存储费用, 但却增加了扫描、分析的计算量, 因此未能从根本上克服离散网格模型时空耗费大的弱点。

拓扑层次模型以计算机图形学中的图形数据结构为基础, 由点表示线、线表示面的数据结构以及记录各活动空间相互位置拓扑关系的存储单元所构成。这种模型与离散网格模型相比, 节省了存储费用, 空间之间的相互距离、面积等几何属性计算较方便。其缺点是拓扑关系属性是由各活动空间的位置坐标通过数学计算来获取的, 一旦某一物体的位置、大小、形状发生变化, 就需要搜索所有的关系记录单元, 重新计算相应属性, 计算繁琐。

基于位置坐标的这类模型的主要优点是: 表达直观、简单, 直接与最终布局图相对应; 便于表示复杂形状活动空间。其主要缺点是存储费用与计算分析费用较大和不便于表示抽象的空间位置关系约束。

针对基于坐标模型的不足, 一些研究者提出了以活动空间的拓扑关系为基础模型, 即拓扑关系模型。这类模型较典型的有图论模型和墙模型。

1964 年 Levin 提出了基于图论的相邻图(adjacency graph), 将其用于建筑平面设计。相邻图表达空间布局的基本规则是: 用图中的结点代表各活动空间或布局区域的边界, 用连接结点的边表示活动空间之间可能的相邻关系。继 Levin 之后, 许多人将相邻图模型应用于厂区设备布局、建筑房间布局等领域。用图论方法解决问题一般分为找到满足某些拓扑约束的较优的相邻子图和从相邻图导出布局图(layout) 等两个子问题。第一个子问题被抽象为最大平面图问题(maximal planar graph) 来加以研究, 如 Foulds 将问题抽象为 GTM。图论模型(Graph Theoretic Model, GTM) 的给定加权图 $G = (V, E)$, 找出一个平面 $G = (V, E_1)$, $E_1 \subseteq E$, 使得 G 总边权值最大。

事实上, GTM 问题是 NP-完全的, 由于问题不可能找到多项式时间算法, 研究者陆续提出 GTM 问题的一些启发式搜索方法, 较典型的是 Foulds 等人 1978 年提出 deltahedron 方法, 以及 Eades 等人 1982 年提出的轮扩展(wheel expansion) 方法以及贪心(greed) 方法。根据具体问题的目标函数, 提出有效的启发式算法, 但力求以较高效率解决相邻图子图的优化构造问题仍然是研究者所进一步研究的目标。

图论方法的第二个子问题是将相邻图映射或转化到欧氏平面(Euclidean plane)。这涉及到增加空间尺寸、选择满意或优化的结果。一些研究者已提出了解决此问题的一些方法。针对相邻平面图只能表达相邻关系的缺点, Flemming 提出了对角顶点图模型, 晏敏等人提出了方图模型。这些模型能够表达方向关系, 适合于解决有方向约束的建筑平面布局问题。此外, 在集成电路布局研究中, 也使用了图论的方法。另一个简化的算法是

Roth 1982 年提出的,该算法与 Mitchell 方法的不同之处在于 Roth 将平面三角图分解成 x 方向和 y 方向两个子图,把二维问题分解成两个一维问题,分别求出 x, y 方向的尺寸,最后据所得尺寸绘出平面布局图。这一方法相对减少了计算量,但往往导致物体相互重叠的布局。

将无尺寸平面布局转化为平面布局坐标图,往往可归结为非线性规划问题;需要考虑包括面积、比例、成本等多因素时,可归结为多目标优化问题。基于相邻图的图论方法主要问题是,一个相邻图可以有多个同态图,因此可转化为多个三角图。在对三角图的偶图着色,或将其分解为 x, y 方向两个子图的过程中,也有许多组合方案,因此在将相邻图转化为无尺寸布局时,仍存在组合爆炸问题。其根本原因是这一途径只能利用“相邻”这一限制关系来剪切搜索分支,未能利用许多如方向、空间大小等其它约束知识。

墙模型(wall representation)是由耶鲁(Yale)大学 Flemming 在 1978 ~ 1980 年提出。1990 年晏敏等人对其进行了进一步的研究,构造了基于墙模型的平面布局专家系统框架。墙模型的基本思想是活动空间是由墙围成的,因此可将活动空间的布置转化为墙的构造过程。每一 w 用一个三元组 $t(w) = (r, A, B)$ 表达,其中 $A = (a_1, \dots, a_m)$ 是在墙 w 的上(或左)边与墙 w 邻接的活动空间有序集; $B = (b_1, \dots, b_n)$ 则是从墙 w 的下(或右)边与墙 w 邻接的活动空间的有序集。 $r = 1$ 时表示水平墙; $r = -1$ 时表示垂直墙。基于墙模型的布局过程分为两个步骤:第一步,由墙生成规则 R 生成一些满足相邻关系的墙表达式;第二步用数学规划求尺寸。墙模型具有几何意义直观、易于计算机表达、可方便地从模型中提取尺寸限制、可自动构成数学规划方程等特点。其缺点也是只能利用“相邻”这一空间关系,同样难免组合爆炸问题;另一缺点是该模型只能处理空间的紧密布局问题。

(3) 空间布局问题的求解 空间布局问题求解算法可分为两类:一类是模型基算法,如在离散网格模型中对二维数组检索、扫描,以确定空间的距离及位置关系的算法,在相邻图模型中构造三角图及其偶图的算法,在八叉树模型中物体干涉的检测算法等皆属此类。另一类是与状态空间的搜索或优化有关的算法,这类算法在设计求解中往往具有通用性,它一般分为(非)线性规划、动态规划等数学规划方法,基于状态空间的搜索方法、约束满足技术,基于知识的方法,基于事例的推理方法以及形状文法方法等。

在布局设计求解中,线性或非线性规划、动态规划等数学规划算法主要用于一维布局或物体位置关系已经基本确定的情况下,进一步求精布局的尺寸。数学规划方法在拓扑、几何分阶段方法中应用较为广泛。在基于状态空间的搜索方法中,设计被视为对状态空间的搜索过程。对状态空间的搜索涉及基于目标和约束而做的决策。在各种常用的布局搜索方法中,根据状态空间结构及搜索策略的不同,可分为树搜索、启发式搜索和局部搜索等方法。

树搜索方法(tree-search)将布局设计求解视为按着一定的顺序确定一些决策的过程。每一步决策涉及几个选择,由此形成树的分枝。在无约束知识的情况下,树搜索方法将陷入组合爆炸。常用的搜索控制策略是分枝界限法(branch and bound)与最佳优先搜索法(best-first search)。由于树搜索方法求解过程容易控制,因此被广泛应用于具有定性或离散特点的问题求解过程。在树搜索方法中主要需解决的问题是:如何确定决策变量及其顺序;如何利用启发信息剪切(prune)搜索分枝,缩小搜索空间。

启发式搜索(heuristic search)在搜索中根据问题的当前状态及约束信息选择搜索的方向或范围。由于穷举式搜索必然导致组合爆炸,因此几乎所有的布局设计算法都采用了与问题相关的各种启发式搜索方法。其中在布局中用到的典型启发式方法是“核心生长法”、基于几何特征的方法、贪心法等。

局部搜索(local search)及随机搜索(random search)方法主要用于布局方案的后期进一步改进。该方案通过局部改变某一或某几个设计变量的值,看结果是否有改进,若有改进则接受改进的方案,如此反复叠代,直到结果满意为止。在布局设计中典型的局部搜索方法有 Shaviv, Francis, Liggett 等人的活动空间成对交换方法;Hiller 的“p-step”方法及其改进方法。随机搜索方法有模拟退火方法等。

在同一问题中,根据不同阶段可分别采用不同的搜索策略,或将两个以上策略综合用于同一设计阶段。一般来说,在约束布局问题中多用树搜索及启发式搜索方法。搜索方法是一种基本的问题求解方法,几乎所有布局设计系统都或多或少地用到了此方法。搜索法往往与约束满足、推理等其它方法相结合。

约束满足技术(constraint satisfaction techniques)其设计结果应该满足多种约束(即设计要求与限制可被看成对变量的约束,而设计的结果应满足所有的约束条件),因此,约束满足技术成为设计求解的研究热点之一。这一方法可分为两个步骤:首先对问题进行描述;然后提供描述求解结果的变量、相应取值域以及变量间的约束。搜索算法就是确定变量取值的一致集合。约束满足技术给出了约束情况下设计空间的清晰描述,其求解过程容易控制,可被用于多约束的设计问题中,研究者在布局设计中也采用了这一技术。

基于知识的方法(knowledge-based system)是在约束布局设计的研究中使用较多的一种启发式方法,也即专家系统方法。专家系统方法用于设计的主要问题在于知识的提炼、获取非常困难,成为其瓶颈。由于知识在设计中的重要作用,基于知识方法的研究与设计自动化的进程是紧密相关的。

基于事例的推理(case-base reasoning, CBR)方法是 AI 的一个重要研究方向,现已开始用于 ICAD 领域。在建筑布局的研究中已有人提出了 CBR 的途径, CBR 方法用于设计系统力图克服设计生成与搜索的复杂性,并试图绕过缺乏领域深层理论之问题,其根本在于克服知识基系统知识获取困难的问题。CBR 途径目前尚未真正克服知识获取瓶颈问题的主要原因:其一是事例的表达、检索、改进均依赖于所用的模型,不能保证所构造的模型能够嵌入事例中所隐含的所有知识;其二是事例一般只是设计的最终结果,并未反映关于设计过程的知识;其三是事例检索匹配及改进需要知识引导。因而,由 CBR 途径能否设计出实际的智能布局系统,尚待进一步研究。

形状文法方法(shape grammar)是 1980 年由 Stiny 提出的。形状文法是一种产生形状语言的形式化重写系统(formal rewrite system)。形状(shapes)是指空间图元(包括点、线、面、体或超平面)的有限布置(finite arrangement)。形状文法包括:终结及非终结图元的集合、初始形状以及形状重写规则的集合等几个部分,其核心是利用形状重写规则来生成复杂图形。Stiny、Gips 以及 Krishnamurti 等人对形状文法在建筑设计方面的应用进行了大量的研究。形状文法方法除了用于建筑形体设计之外,目前已应用于图案创作等领域,如浙江大学研制的“智能模拟彩色平面图案创作系统”。由于形状文法方法难

以反映布局设计中的空间相互关系等概念,系统难以抽取形状所代表的语义,因此其推理常带有盲目性,还存在构造形状文法困难以及状态空间不完备等缺点。

5 3 2 三维正交方位图模型及其布局搜索算法

人们在处理复杂对象时,往往是先将问题简化。同样研究空间布局问题,首先也应建立适合于描述问题的抽象模型结构,通过对三维空间关系的分析,借鉴以往二维空间建模的空间思想,构造一个适合描述三维空间关系的三维正交方位图模型。然后通过引入正则约束概念,实现基于三维正交方位图的状态空间布局的搜索算法。其布局设计过程如图 5 .12 所示。

图 5 .12 布局设计求解过程

1 . 三维正交方位图模型

为了构造一种有效的三维空间布局模型,需先从布置物间的位置关系入手。为了简化问题,可把布置物抽象为一个长方体(一个复杂形体可视为长方体的组合),同样整个布局的空间区域也可视为长方体形状。

用三向图可以表示空间布局,其规则为:(a) 三向图的外点 AE, BW, FT, BK, LT, RT 可分别表示布局的包围长方体的上、下、前、后、左、右六个面;(b) 内点表示待布置的长方体;(c) 若两个长方体存在某种方位关系,则应在图中增加相应边(ab 表示自上而下, fb 表示自后而前, lr 表示自左而右);(d) 用各内点与六个外点分别连接固定的有向边来表示待布置长方体与包围长方体相应面的固定方位关系。

三维正交方位图模型表示布局的规则为:(a) 内点与外点分别表示待布置长方体及包围长方体的各面。(b) 两长方体存在的方位关系(可能同时有多种)总有一种,而且只允许一种可通过图中对应两内点间的相应边或路径加以体现。若两个长方体或者长方体与包围长方体某一面的方位关系已经由图中某两条及两条以上连接边的路径加以表示,则图中该两点间不允许再有直接的连接边。(c) 用内点及外点间固定的路径关系来表示各长方体与包围长方体各面间的方位关系。

同时也必须注意到,尽管两个长方体间可能存在两种以上方位关系,但在同一三维正交方位图中仅允许表示一种,这可能导致同一布局有两个以上的三维正交方位图的表示,但这些表示之间是相容的,它们均体现了布局中的若干关系,允许忽略某些空间关系。应注意到,三维正交方位图是表达不重叠的三维布局的具有最少拓扑关系限制的模型结构。此外,此模型可以表达一定的相邻关系。

2. 基于三维正交方位图的搜索树

基于模型的三维布局方案的生成,可以概括为三个方面:(a) 将约束转化为空间关系的要求;(b) 在模型构造的状态空间中搜索,生成满足特定的空间关系以及其它约束(这些约束通常不容易直接转化为空间关系的要求,但可以在求解过程中进行检测)的状态结点;(c) 将基于模型的状态结点转化为具体坐标图。过程(a)可通过引入专家系统加以实现,过程(c)可由尺寸推理计算自动完成,而过程(b)是研究基于三维正交方位图模型的搜索生成算法。为区别三维正交方位图中的结点,我们将状态空间搜索树中的结点,称为状态结点。

粗略地讲,生成满足约束的三维正交方位图可以是在满足约束的前提下,不断增加结点的过程。若 S_K 为含 K 个内点的三维正交方位图,则对于 n 个单元的布局问题,其过程可描述为:

$$S_0 \quad S_1 \quad S_2 \quad \dots \quad S_n$$

若当前状态为 S_{i-1} ($1 \leq i \leq n$), 并且下一步骤应选择某一位置单元 u_i , 将 u_i 连接到 S_{i-1} 中生成满足约束的三维正交方位图 S_i , 这一过程可称 i -connect 过程。若在某一 connect 过程时失败,则应回溯到前一状态,若 1-connect 过程(即从 S_0 到 S_1)失败,则无解,若 n -connect 过程成功,则求解成功。

3. 基于 AHP 的启发函数构造法

在人工智能问题求解中,需要研究如何利用有关问题域的启发信息引导状态空间的搜索,以避免组合爆炸,降低计算费用。在单纯采用专家系统技术求解设计问题时,由于知识的不完备性,在搜索过程中往往不能确定下一步应当扩展哪一个节点,从而造成盲目搜索。在实际设计中存在约束不足问题,解空间内含有很多满足基本约束的解,由于搜索能力的限制,只能搜索到其中一部分解,因此对搜索的控制策略将影响到实际解的生成。设计往往涉及多准则决策,而在约束不足情况下由盲目搜索求得的满足基本约束的解,一般在多准则下不是较优的。因此在设计求解中,按照某一评估标准,即启发函数来确定下一步应扩展的节点,采用最佳优先搜索,是避免盲目搜索造成的组合爆炸,并在有限的时空界限内求得较优解的重要策略。

在研究计算机结构布局设计的过程中,针对多准则布局设计问题,人们提出了一种启发函数 AHP 的构造方法。该方法对工程设计中启发函数的构造具有一定的指导意义。此方法可推广到一般设计问题,为在设计求解中构造启发函数提供了一种有用的方法。

(1) AHP 方法简介 AHP(analytic hierarchy process)方法,是一种多准则、定性与定量相结合的决策分析方法。AHP 方法解决问题,大致分为四个步骤:首先是建立递阶层次结构。把系统中所需考虑的各因素或问题分组,形成不同层次,一个层次的元素作为准则,对下一层次的某些元素起支配作用,同时又受上一层元素的支配,这样构成问题的递阶层次结构,从上到下各层次分别为目标层、准则层(可能有多层)、方案层(表示决策方案)。第二是构造两两比较矩阵。假定上一层的元素 C_k 作为准则,对下一层次元素 M_1, M_2, \dots, M_n 有支配关系,为了得到 M_1, M_2, \dots, M_n 对于 C_k 的相对重要性的权重,

AHP 采用了两两比较的方法,即针对准则 C_k ,两元素 M_i, M_j 哪一个更重要一些,并对重要多少赋予一定数值,由此得到两两比较判断矩阵 $A = (a_{ij})_{nn}$ 。第三通过解 $A_w = \max W$,得到 A 的特征根 w ,将 w 正规化后,即可作为元素 M_1, M_2, \dots, M_n 在准则 C_k 下的排序权重。最后,自上而下逐层计算每一层次中所有元素相对于总目标的相对权重,最终得出最低层元素,即决策方案层诸元素对于总目标的相对权重。

(2) 启发函数的构造 进行最佳优先搜索之关键是建立合适的启发函数,启发函数是对费用的估算,即设计费用是对设计优劣的综合评价结果,包括对功能、造价等标准的综合评估。若用 $f(n)$ 表示节点 n 的启发函数值, $g(n)$ 表示从起始节点到节点 n 的最小费用估算, $h(n)$ 表示从节点 n 到目标节点的最小费用估算,则有 $f(n) = g(n) + h(n)$ 。

在设计问题中,费用(指对设计优劣的综合评价结果)的内涵通常涉及造价、功能、性能及维护等多方面,是诸费用的综合,因此,费用估算属于多准则决策问题。通过运用 AHP 方法建立递阶层次结构,分析计算方案层对目标层(即费用)的影响权重,然后据此构造启发函数,这就使得在设计求解中可以运用最佳优先搜索策略。

在设计型问题中,费用通常没有准确的希望量度,且由于计算完整信息的工作量很大,因而应选择典型信息来构造启发函数 f ,以便能通过适量搜索找到好的(但不一定最优)结果。构造 f 的基本原则是:使 f 的计算量(时间)尽可能少,并且 f 应尽可能体现主要的一些费用,这两方面相互制约,一般要在这两者间进行折中。下面,以计算机结构布局设计为例来说明构造启发函数的具体方法。

在布局设计中,综合费用可以视为布局方案的函数,而方案可以由待布置单元及其相互关系加以描述,因此费用主要取决于待布置单元及其关系。如果运用 AHP 方法建立递阶层次结构,则目标层应为综合费用,方案层应涉及一些重要的空间关系。由于可能空间关系的组合太多,不宜引入过多的空间关系,以免片面追求信息完整而导致计算复杂度增加;另一方面复杂空间关系一般可由两单元间的相邻与方向关系体现。

三维正交方位图模型实质上可视为相邻与方向关系的一种集成表示。基于该模型研究布局方案对费用的影响,应将方案中最关键的要素,即两单元间的相邻与方向关系作为方案层的基本元素,通过 AHP 方法,求得它们对综合费用的排序权重,然后通过这些排序权重,进一步去构造与计算启发函数。在计算机结构布局设计中,采用了在计算机结构布局中空间关系对费用的影响之递阶层次结构(图 5.13 所示),对计算机结构空间布局的讨论也可推广到一般设计问题。根据图 5.13 的递阶层次结构,运用 AHP 方法可以求得方案层各元素对综合费用的排序 $w_{\text{near}}(x, y)$, $w_{\text{direct}}(x, y, r)$, area-weight (其中 $w_{\text{near}}(x, y)$ 表示两布置物 x 与 y 相邻之权重, $w_{\text{direct}}(x, y, r)$ 表示 x 在 y 的 r 方向之权重, area-weight 表示布置区域表面积之权重)。

在进行设计求解(即对设计状态空间搜索)之前,首先计算排序权重,将其存储起来,以便计算启发函数值。在设计求解中,与设计方案有关的要素不断添加到设计方案中,而这些要素将影响着设计的费用。由于已用 AHP 方法得出了这些要素对费用影响的权重,因此,容易据此分析当前状态下的费用,并作出估计。在计算机结构布局设计实例中,根据这一思想,具体构造 $g(n)$ 如下:

$$g(n) = \sum_{x,y} (w_{\text{near}}(x,y) / (d(x,y) + 1)) + \alpha \cdot w_{\text{direct}}(x,y,r) + \beta \cdot \text{surface-area} \cdot \text{area-weight}$$

式中, x, y 表示任意两布置物, $d(x, y)$ 表示 x, y 间的距离, $w_{\text{near}}(x, y)$ 表示 x 与 y 相邻对费用影响之权值, $w_{\text{direct}}(x, y, r)$ 表示 x 在 y 的 r 方向对费用影响之权值, r 是节点为 n 时 (即当前状态下) x 对于 y 的实际方向。若该方向尚未确定, 则 $d(x, y)$ 与 $w_{\text{direct}}(x, y, r)$ 均取零值, surface-area 表示节点为 n 时布局区域最小长方体的表面积, 其中 α, β 为正数, 用于调整 $g(n)$ 中三个加数对 $g(n)$ 的贡献, 一般情况下, 可取 $\alpha = \beta = 1$ 。

图 5 .13 递阶层次结构

从当前状态到设计目标状态的费用估计主要取决于当前状态下尚未确定的方案要素的情况, 在本例中, 规定 $h(n)$ 为:

$$h(n) = m^2 + l$$

式中, m 为除已置单元和当前待置单元外的未置单元的个数, l 为当前待置单元与已置单元中尚未确立关系的单元数。在这里构造 $h(n)$ 应体现未确定的设计参数愈少与设计目标愈接近 (因而进一步的 $h(n)$ 愈小) 的思想。

最后采用 g 和 h 的相对加权 $f(n) = g(n) + w \cdot h(n)$ (其中 $w > 0$)。在状态空间搜索的过程中, 通过对节点计算其 $f(n)$ 值, 决定下一步优先扩展的节点。由于各方案要素对费用的影响是一次性静态计算后存储起来的, 因此在搜索求解时, 计算启发函数仅是对存储结果以及现有设计情况的简单查询计算, 其计算量很小。该方法可避免设计求解中的组合爆炸问题, 有效地降低计算复杂度。

以上对计算机结构空间布局的讨论也可推广到一般设计问题, 由此得到设计型问题求解的最佳优先算法如下:

根据具体设计问题, 将影响设计费用的诸因素 e_1, e_2, \dots, e_n 按照其支配关系分解为递阶层次结构, 其最上层为设计费用, 其最下层为方案层, 其元素设为 M_1, M_2, \dots, M_r 。

为简化分析,各层次应体现主要因素,忽略次要因素。

根据实际问题的经验,对逐层中每两个元素相对于上一层某固定元素赋以两两比较重要性数值。

根据 AHP 算法,计算方案层 M_1, M_2, \dots, M_r 对费用影响的相对权重 w_1, w_2, \dots, w_r , 将其存储起来。

根据设计状态空间模型与以上权重构造启发函数。在状态节点为 n 时的费用估计 $g(n)$ 一般可表示为 $g(n) = \sum v_i(n) \cdot w_i$, 其中 $v_i(n)$ 为在状态 n 时与方案 M_i 相关的某一函数。若在状态为 n 时,方案 M_i 尚未确定,则 $v_i(n)$ 取 0 值,否则根据具体问题取某一正值,然后根据中间状态中尚未确定的设计参数确定当前状态 n 到目标状态的期望费用 $h(n)$ 。最后由两者之和确定启发函数 $f(n) = g(n) + \alpha \cdot h(n)$ (其中 $\alpha > 0$)。

在状态空间搜索过程中,对搜索到的节点根据启发函数公式计算其启发函数值。每次总是先扩展 OPEN 表中启发函数值最小的节点,实施最佳优先搜索,直到搜索到所有设计参数均完全确定的满足约束的目标状态为止。

5.3.3 基于正交方位图与遗传算法的包装布局

1. 包装与组合优化

包装(如商品包装)布局问题,可描述为将已知形状、大小的若干个物体包装在一起,使其包装体积最小或包装材料最省。为简化讨论起见,这里被包装的物体形状以及外包装的形状均假设为长方体。这种假设通常符合大部分的情况。另一类与此相近的问题是外包装箱子的形状、大小已知,要求将尽可能多的被包装物体装入箱子内,如集装箱货物装载问题就属于此类。以上两类布局优化问题有相近的性质,实现方法可相互借鉴,故在此仅对第一类问题加以讨论。

如果抽去布局问题的尺寸连续性特征,则可将其视为组合优化(combination optimization)问题。所谓组合优化是指在离散的、有限的数学结构上寻找一个满足给定约束条件,并使目标函数值达到最大或最小的解。一般来说,组合优化问题通常带有大量的局部极值点,往往是不可微、不连续、多维的、有约束条件的、高度非线性的 NP 完全问题,因此精确地求解组合优化问题的全局最优解,一般是不可能的。遗传算法作为一种新型的模拟生物进化过程的随机化搜索、优化方法,近十几年来在组合优化问题中得到了广泛的应用,在解决诸多典型组合优化问题中显示了良好的性能和效果。然而,G.D.Smith 指出,尽管已有很多用遗传算法解决优化问题的例子,但是用遗传算法解决实际问题的例子却还非常少。Nissen 指出,在实际问题中,还未看到进化算法的重大突破。Radcliffe (1992 年)和 Kathryn(1996 年)等指出,目前遗传算法还难以有效处理实际应用问题中的复杂约束(highly constrained problems)。1994 年,Michalewicz 指出,遗传算法的罚函数法(penalty function approaches)对于复杂约束问题运行效果欠佳。1996 年 Kathryn 指出,针对问题的不同领域,将经典遗传算法框架与问题知识以及其它启发式策略杂交可能是解决实际约束问题,提高求解质量的可行途径。因此,在实际问题中,构造与问题结构知识有关的编码及交叉、变异算子,是用遗传算法解决实际约束问题的主要途径。

2. 包装布局问题的编码设计

(1) 编码原则 遗传算法主要是通过遗传操作对群体中具有某种结构形式的个体施加结构重组处理,从而不断地搜索出群体中个体间的结构相似形,形成并优化积木块以逐渐逼近最优解。编码的策略或方法对于遗传操作,尤其是对于交叉操作的功能有很大影响。在很多情况下,编码形式也就决定了交叉操作。编码问题往往称之为“编码-交叉”问题。由于在复杂的实际问题中往往涉及到约束问题,因此“编码-交叉”设计非常困难,这是遗传算法问题能否用于实际问题的关键。编码(coding)是问题空间向 GA 空间的映射,由 GA 空间向问题空间的映射称为译码(decoding)。编码策略通常采用 De Jong 的编码规则:有意义的积木块编码法则,即所定编码应当易于生成与所求问题相关的短距和低阶的积木块;最小字符集编码规则,即所定编码应采用最小字符集,以使问题得到自然的表示或描述。

由于布局问题方案涉及到布置物间的拓扑关系,而这些拓扑关系又具有特定的制约关系,从这个意义上说,布局问题属于约束问题,故其编码设计问题除兼顾以上规则外,应主要考虑交叉操作中避免生成致死基因的编码。

(2) 模型分析与变换 根据以上讨论可知,正交方位图模型是一种有效的表达不重叠布局的模型。实际上,该模型中规定的空间位置关系是表达不重叠布局的最小的拓扑要求。因为在包装问题中,目标是使包装体积或面积最小,而拓扑约束的增加将可能导致包装体积或面积过大,因此在保证布局中物体不重叠的前提下,希望对布局本身施加最少的约束。在这一意义上,对于包装问题来讲,正交方位图显然是理想的布局模型。这是选择正交方位图作为包装布局模型基础的主要原因。

正交方位图模型可以由矩阵来加以表达。根据 x, y, z 三个坐标方向,将其分解为三个子图,分别对应三个子矩阵,每个子矩阵的元素为 0 或 1。若在与某一方向对应的子矩阵中,第 i 行第 j 列元素为 0,则表示从 i 到 j 在该方向无方位关系,1 则表示在该方向有方位关系。

3. 基于知识的布局遗传算子

遗传算法由于仅靠适应度来评估和引导搜索,所以求解问题所固有的约束条件不能明确表示出来。在实际应用中,多数问题是带有约束条件的,如货郎担问题(TSP)等。布局问题从编码到约束条件等方面比 TSP 问题复杂得多,因此用遗传算法解决布局问题,必须考虑一些对策。

似乎可以采用一种十分自然的方法来考虑约束条件,即在进化过程中,迭代一次就设法检验一次新的个体是否违背了约束条件。如果没有违背,则作为有效个体,反之,作为无效个体被去除,这种处理方法对于弱约束问题求解还是有效的,但对于强约束问题,求解效果难以保证。因为在这种场合,寻找一个有效个体的难度不亚于寻找最优个体。解决带约束问题的一种策略是采取惩罚方法(penalty method),即将个体违背约束条件的情况给予惩罚,并将此惩罚体现在适应度函数中,这就将一个约束优化问题转换为一个附带考虑代价(cost)或惩罚(penalty)的无约束优化问题。惩罚函数法原则上也只适用于弱

约束问题,因为该方法没有从遗传算法本身入手考虑如何得到有效个体,因此它本身不能保证遗传算法以较大概率来生成有效个体。

遗传算法本质上属于搜索算法,Adrian A .Hopgood 曾形象地把搜索问题比喻为“猴子弹钢琴问题”,即设想在一架巨大钢琴上,猴子随机地弹奏琴键,最终能够弹奏出贝多芬的某一旋律来,这几乎是不可能的。原因在于搜索的空间是如此巨大,在实际可用的时间框架之内,猴子根本不可能撞上这种旋律;只有具有音乐知识的演奏者,才有希望产生这种旋律,因为他可以用知识去纠正搜索范围与方向。因此要解决实际布局问题,必须从遗传操作——主要是交叉操作的设计上采取必要措施,运用布局问题的领域知识以生成有意义的个体,即采用基于知识的启发式交叉方法。

(1) 基本交叉算子 由于在布局编码个体的基因之间存在约束制约关系,交叉生成子串后出现矛盾或冗余,而导致不好甚至致死基因的概率一般来说是较大的,因此必须基于知识对其修正。为了保证修正后的个体能够较多地保持父串的特征,将串长度较大的一部分固定下来,对另一部分进行修正。修正时首先用固定部分的非零基因屏蔽非固定部分的等位基因与反射基因(基本交叉算子步骤),然后逐一在非固定部分考虑添加未被屏蔽掉的基因(步骤),最后考虑应使每一组等位基因与反射基因中恰有一个不为零,若全为零则应修改之。

基于方位图知识的基本交叉算子按以下步骤构造后代:

设父串 $S = a_1 a_2 \dots a_{n2} b_1 \dots b_{n2} c_1 \dots c_{n2} d_1 \dots d_n = x_1 \dots x_i \dots x_{3n+n}^2$ 以及 $S = a_1 a_2 \dots a_{n2} b_1 \dots b_{n2} c_1 \dots c_{n2} d_1 \dots d_n = x_1 \dots x_i \dots x_{3n+n}^2$ 均为有效个体。

随机设定交叉处(采用一交叉点),交叉点 S 及 S 各分成两部分 $S = S_1 | S_2$, $S = S_1 | S_2$ 。根据交叉点在串中偏左或偏右,以决定哪一部分为主(其基因不变),哪一部分为从(其基因要根据主部分调整)。若交叉点偏左,则右边部分为主;反之,左边部分为主。

假定左边部分为主(不失一般性),则子串后选个体为 $N_1 = S_1 | S_2$, $N_2 = S_1 | S_2$ 。因 $S_1 \subseteq S$, $S_2 \subseteq S$,故 S_1 以及 S_2 均无矛盾和冗余。

若 S_1 的长度大于等于 $3n^2 - 2n$,则 N_1 与 N_2 即为交叉所产生的有效个体,转 。否则,将 S_2 分为两部分 $T_2 | U_2$,其中 U_2 为放置方式编码;同样处理 $S_2 = T_2 | U_2$ 。

固定 S_1 ,对于 S_1 中每一非 0 基因,检查 T_2 中相应的等位基因及反射基因,如有等位基因及反射基因则将其改为 0 值;同样固定 S_1 对 T_2 进行类似处理得到 V_2 ,这样得到 $N_1 = S_1 | V_2 | U_2$, $N_2 = S_1 | V_2 | U_2$ 。

令 W_2 是与 V_2 等长的二值编码,初始令其编码全为零。任取 V_2 中某一非零值(假定为 $V_{Z(i,j)}$, $V_{Z(i,j)}$ 是指第三个子阵 Z 的 (i,j) 位基因)。在 W_2 中相应位置的编码 $W_{Z(i,j)}$ 设为 1,此时由增加 $W_{Z(i,j)} = 1$,可能导致 Z 中一系列传递关系,处理这些关系的方法为:

$$k \in \{1, \dots, n\}, Z(k, i) = 1 \rightarrow Z(k, j) = 1; Z(j, k) = 1 \rightarrow Z(i, k) = 1$$

检查由于增加 $W_{Z(i,j)}$,是否导致与 S_1 部分矛盾或出现关系冗余,或与 W_2 中已增加过的关系出现矛盾或冗余,若是,则取消该编码值及相应的传递关系;若否,则固定该次操作。在 V_2 中标记该位置 (i,j) ,以上过程进行到 V_2 中所有非 0 值全部标记完为止,最后由 W_2 得到 X_2 ,综合得到 $M_1 = S_1 | X_2 | U_2$,同理可得到 $M_2 = S_1 | X_2 | U_2$ 。

检查 M_1 中的三矩阵 X_1, Y_1, Z_1 , 对于 (i, j) , 检查 (i, j) 与 (j, i) 位是否全为 0, 若是, 随机增添某关系, 即在 X_1, Y_1, Z_1 中任选一个矩阵的 (i, j) 位或 (j, i) 位赋值 1, 然后如同 中所做的工作, 将由其产生的传递关系增加于矩阵中, 若某传递关系与原关系发生冲突(矛盾或冗余), 则放弃该选择, 另选其它, 直到合适为止。检查完所有的 (i, j) (其中 $i < j$), 即生成一个合法有效基因 O 。同理检查 M 得到 O_2 。串 O, O_2 即为交叉所得的有效个体。

结束。

在基本交叉算子作用下, 子串能够部分或大部分地继承父串的结构特征和有效基因。

(2) 辅助交叉算子 前述交叉算子能够较好地保持方位关系和积木块繁殖, 全局搜索的能力较强, 然而在布局方案编码中除了方位关系外, 还有某种邻近关系, 而这种邻近关系在上述基本交叉算子中不能得到有效繁殖。因此当希望适当保持这种邻近关系时, 可使用以下基于邻近知识的交叉算子。

(3) 布局限制交换算子 交换是一种模拟生物体基因换位变异的遗传操作, 相对于交叉算子而言, 交换操作属于基因的局部操作, 属于局部搜索的范畴。在布局设计中, 一种典型的局部搜索方法是所谓成对交换方法, 通常是交换两个布置物的相互位置, 而其它布置物位置不变, 在无(弱)约束布局问题中, 这种方法常被采用。根据布局问题的这一特点, 引入一种限制交换算子。

该算子操作由以下各步组成:

选择个体 S ;

随机选择所涉及的物体 i, j (其中 $1 \leq i, j \leq n, i \neq j$);

个体串 S 按前面所述可分为四部分 X, Y, Z, W , 对于 X , 将其第 i, j 行互换, 然后将 j, i 列互换, 得 X' , 同样处理 Y 与 Z , 得到 Y', Z' 。

X', Y', Z', W 即为限制交换后的新个体。

4. 基于知识的初始群体生成

在基本遗传算法中, 初始群体中个体的每位基因均是随机独立地产生的, 显然在约束问题求解中采用这种随机方式生成的初始个体不符合约束的可能性很大。如果初始个体是非合法的, 那么经过交叉产生的子代合法的可能性也很小, 在布局问题中这种方式是不适用的, 本章构造交叉算子及交换算子时有个前提就是父串是合法的。这自然要求初始群体中的个体是合法的, 故需要研究如何生成合法个体, 并保持群体中个体的多样性。因此, 在构造初始群体时, 既要随机的选择基因的取值, 又要根据布局中基因编码间的内在关系知识引导基因值的确定和选取, 此处需要注意由编码关系的传递性导出的编码值和避免等位基因与反射基因的取值冗余这两点。基于以上考虑, 初始个体按以下规划产生:

For ($i = 1; i \leq n; i++$)

For ($j = i + 1; j \leq n; j++$)

{

(a) 检查 (i, j) 的所有等位基因与反射基因, 是否已有某一非 0 值, 若是, 转结束;

(b) 在 (i, j) 的所有 6 个等位基因与反射基因中随机选择一个, 赋值为 1; 不失一般性, 假定选中 X 矩阵的 (i, j) 位置;

(c) 在矩阵 X 中作如下处理:

For ($k = 1; k \leq n; k++$)

{

if $X(k, i) = 1$

若除 X 中的 (k, j) 位外, 在 X, Y, Z 中 (k, j) 位与 (j, k) 位全为 0, 则令

$X(k, j) = 1$, 否则返回 (b) 重新选择;

endif

if $X(j, k) = 1$

若除 X 中的 (k, j) 位外, 在 X, Y, Z 中 (i, k) 位与 (k, i) 位全为 0, 则令

$X(i, k) = 1$, 否则返回 (b) 重选其它;

endif

}

}

For ($i = 1; i \leq n; i++$) 随机从 $\{0, \dots, 5\}$ 中选择布置物 i 的布置方式。

5. 合适值的计算

在此以商品包装设计为例来说明包装布局问题。通常关心的是包装设计的总体积和包装用料的多少, 这两方面均涉及到计算外包装长方体的尺寸。由于个体编码是基于诱导方位图的, 而诱导方位图中无环路, 因此可以将诱导方位图视为赋权图 (据布置物长方体的边长赋权)。这样外长方体在某一方向的尺寸问题就可化为最小关键路径长度问题, 可采用 E. W. Dijkstra 于 1959 年提出的 Dijkstra 算法求解, 或采用针对诱导方位图而提出的如下算法 (注意此处图的无环路假设)。对于诱导方位图, 采取这两种算法, 其复杂度上界均为 $O(n^2)$ 。

最小外包装长方体尺寸计算方法:

设基于诱导方位图 G 的个体编码 L , G 的子图为 G_x, G_y, G_z , 以下据 G_x 来计算最小外包装长方体的 x 方向尺寸;

设 $G_x = (V, E)$, 令 $h(v)$ 为物体在 x 方向的自身尺寸, 初始时 $f(v) = h(v), v = v_1, \dots, v_n$;

For 每一 $v = v_i, i = 1, \dots, n$,

if $(v_i, v_j) \in E$

Then $f(v_j) = f(v_i) + f(v_j) (j = 1, 2, \dots, n)$;

endif

$f(v_i)$ 中最大者即为所求。

为简单计, 此处合适度函数为布局中外长方体的体积的倒数。

针对商品包装布局设计问题, 采用了布局编码及交叉算子、限制交换算子等, 结合布局空间关系模型知识, 构造适当的编码与遗传算子, 运用遗传算法求解复杂的布局优化问题, 采用上述算法对商品包装布局设计是可行的。对于一般规模的布局问题, 该算法能够较快地收敛到最优或接近最优解。

5 3 4 基于约束满足的多策略布局设计算法及其应用

约束满足与搜索、推理一起被成为人工智能问题求解的三大方法。约束满足技术 (constraint satisfactory technique) 在约束求解问题中占有举足轻重的位置, 许多诸如规划、设计等人工智能问题可被编码为约束满足问题 (constraint satisfaction problem, CSP)。CSP 问题就是要在一个变量集合中寻找所有变量一致的取值, 使得这一取值满足这些变量之间的给定约束。如果将布置物间的位置关系视为变量, 布局问题就可以视为一个 CSP 问题。以往 CSP 求解严格区分变量与变量间的约束, 约束不能主动参与求解过程, 不能有效利用约束所含控制信息最大程度地缩小状态空间的搜索范围。可将约束视为变量, 我们将这种变量称之为复合变量, 在搜索中优先考虑。采用这一思想, 可以使约束主动参与搜索控制, 避免“生成测试”方法在搜索中的盲目性, 从而降低计算复杂度。

在实际约束设计问题中, 领域知识是多方面的, 既包括一些给定的静态约束知识, 也包括在设计过程中的推理知识, 还包括由领域启发函数所体现的启发性知识。显然, 仅采用单一的求解策略, 如 CSP 方法、专家系统、启发式搜索等, 均不能完整运用多种类型的设计知识。因此, 研究如何将约束满足技术与搜索及推理技术有机融合起来, 构成一个集成的问题求解环境, 具有重要的理论及实际意义。

1. 设计约束的处理方法

所谓约束, 是指设计变量 x_1, \dots, x_n 间所存在的某种或某些关系 $R_1(x_1, \dots, x_n), \dots, R_m(x_1, \dots, x_n)$ 的限定。如果换一个角度, 将约束本身也视为变量, 而这些关系 R_i 作为其取值, 则约束及其变量就可统一作为设计变量进行处理。

根据约束对变量的限制关系是唯一确定还是可选择的, 可将约束变量分为确定性约束与可选择性约束两种。实际上一个约束是否为确定性约束常取决于所处环境的粒度, 如约束“物体 u 与 v 邻接”在抽象层次上显然应视为确定性约束, 然而在考虑 u 与 v 具体如何邻接时, 显然有几种方案可选择, 此时该约束就应视为可选择约束。在某一环境粒度下, 确定性约束变量取惟一值, 可选择性约束变量可从多个值中取某一值。

若在某一环境粒度条件下 (即设计模型下), 约束可视为设计变量, 那么对约束问题的求解就可以转化为无约束问题求解。在搜索中, 给约束变量以最高优先级优先处理 (树搜索顶层)。将约束尽快融进设计方案中, 缩小搜索空间范围, 可避免“生成测试”所带来的盲目搜索复杂度。当然在许多实际问题中, 有一些约束难以融入设计变量集合, 如布局中的面积要求不能有效地在约束图模型中表达, 但是尽可能将可表述的约束作为设计变量, 将能够有效地缩小搜索空间, 降低问题复杂度。

2. 约束满足与复合变量方法

由于 CSP 模型提供了一种简单、自然并且强有力的知识表达框架, 因而被广泛用于各种与 AI 有关的问题表示与求解, 为真值维护系统 (TMS)、专家系统、约束逻辑程序设计提供了基本的工具。然而, 过去对 CSP 的研究, 局限于基于变量的搜索, 约束所起的作用是对变量取值是否有效进行检查, 约束在 CSP 方法中的角色是被动的。由于约束本身

未能主动参与状态空间搜索过程,从而使得搜索复杂度难以降低,结合基于层次约束图的布局算法的研究,这里介绍一种基于复合变量的约束满足方法。

人工智能问题求解中常涉及到动态环境,约束知识集合并非一成不变,而是要根据要求的不同不断进化的。在实际应用中,当由某一 CSP 所表达的约束知识被用户(交互式问题求解)或其它进程(分布式问题求解)或其它外部干扰(不可预见事件)修改后,就存在如何有效利用已经做过的工作,使新问题尽快收敛到一个解的问题。这就是所谓动态约束满足问题(dynamic constraint satisfaction problems,DCSP)。在对动态约束满足问题所进行的研究中,研究者们局限于对“约束的动态性”,即所谓约束限制(restriction)或者松弛(relaxation)的研究。然而,在实际问题中除约束可变外,变量也可增减。动态性应分别体现在约束及变量(此处即复合变量与基变量)两方面。因此,相对于只允许约束变化的动态 CSP 问题,可把约束及基变量均允许变化的 CSP 问题称为双重动态约束满足问题(dual dynamic constraint satisfaction problem,DDCSP)。

在解决 DCSP 问题时,主要的策略是所谓 Nogood 约束记录方法,该方法通过记录以前搜索出现死端(dead-end)的原因,避免今后相同情况的再次发生。Rina Dechter 把这一过程看作为学习过程。

树搜索方法是求解约束满足问题的基本途径,以往研究者局限于对各基变量的不同取值形成的搜索树所进行的约束满足研究,其本质是“生成测试”的方法,变量赋值的盲目性较大。针对这一缺陷,考察因复合变量(约束)取不同值而形成的搜索树,并借鉴 Thomas Schiex 的 Nogood 记录策略控制搜索。

(1) 在树搜索中对约束取值的学习 在根据约束的不同取值进行树搜索的过程中,可对已经发生的不一致取值进行记录,避免无意义的搜索。举例来说,假设有四个复合变量 G, C, S, A , 其取值域分别为 $d(G) = \{a_1, a_2, a_3\}$, $d(C) = \{c_1, c_2\}$, $d(S) = \{s_1, s_2, s_3\}$, $d(A) = \{a_1, a_2, a_3, a_4\}$ 。按照以上规定的变量及其取值顺序进行树搜索,在搜索中,发现了一些不一致取值的情况。假定这些不一致取值情况是 (c_1, s_2) , (c_1, s_3) , (c_1, a_3) , (c_1, a_4) 。则在今后搜索 C 的同一取值分支 c_1 时,可根据这些不一致取值记录,从 S, A 的取值表中直接去掉 s_2, s_3, a_3, a_4 。由约束取值存在的不一致的情况积累,还可能发现约束矛盾(不一致)。当已确定约束集 C 的某子集为不一致时,即应停止搜索,指出所存在的问题。

(2) 单纯复合变量的 Nogood CSP 算法 当所有约束均可表达为基变量取值的组合限定时,称为单纯复合变量 CSP 问题。此时,可以采用以约束(复合变量)取值为主线的搜索。在搜索中记录已经历过的约束赋值矛盾,可以指导后续搜索过程。

以下为基于约束不一致取值记录,寻找 CSP 问题(X,C)的一致约束取值的回溯算法。

NogoodSearch1(, S, R_s)

\\ 为一组约束的取值,S 为约束集合,R_s 为 S 中约束的取值表

\\ 初始调用时 = , S = C, R_s 为约束初始取值描述

If S =

Then (X,)为一致取值,Stop .

Else


```

Let C = S
For 每一 c ∈ Rc      \\ Rc 为约束 c 当前的取值集合
    RS - {c} = Modify Value( c, NogoodTable)
    \\ 根据 c 及 NogoodTable 修改后续约束的取值表 RS - {c}
    Let      { c }
    [h, P, L] = check( c, ) \\ h 返回值为 0, 1, 2
    \\ P 为约束不一致取值的集合
    \\ L 为不一致约束的集合
    If k = 0  \\ 无约束取值矛盾
        Then NogoodSearch1( , S - {c}, RS - {c} )
    Endif
    If h = 1  \\ 有约束取值矛盾
        Then RecordCompositeNogood( P )
    Endif
    If k = 2  \\ 有约束矛盾
        Return L, Stop .
    Endif
Endfor
Endif

```

其中函数 `check(c,)` 用于检查 C 的取值 c 与 中的其它约束取值是否一致。若有不一致,则记录于 Nogood Table 中,并进一步通过标记确定是否存在两个约束,它们的所有取值组合均有矛盾,若是,则返回这两个矛盾的约束,结束。

3. 多策略 CSP 搜索算法

约束满足技术、基于知识的推理、最佳优先搜索等是布局设计中基本的求解技术。人们在研究中发现,其中每一种方法只反映了设计问题求解的某一侧面。约束满足方法一般仅刻画了设计对象所应满足的某些静态属性。由于在设计问题中,存在不能预先表示的知识,因而,变量及其约束将可能随着设计过程而变化。此外,设计中还存在不能明确表达的约束等问题。布局设计的这些特点,决定了简单的约束满足技术难以有效解决实际设计问题,因此,必须结合其它技术。基于知识的推理方法反映了设计对象依赖于条件的动态可变属性,可以弥补 CSP 方法缺乏动态知识之不足。在实际布局问题中,除了规定约束限制外,往往还要求按照某些准则使设计方案达到最优或较优,即存在一个目标函数。显然,在这类布局问题中,除了应使得设计结果满足约束之外,还应使得目标函数较优。最佳优先搜索的启发式方法反映了设计对象的这种优化特性,一般根据目标函数及问题的结构可以构造启发函数 $f(x)$ 。利用启发函数可以体现某些设计目标、某些设计偏好 (preference) 以及不能明确表达的约束知识。将这三种求解技术有效地结合起来,构成一种综合的求解控制机制,使得各种技术能够共同参与设计过程,这对于完整描述设计对象,限制搜索的状态空间,得到合理的求解结果,是非常必要的。

根据设计理论,设计是一个运用知识从初始产品对象描述向完整产品对象描述的不断迁移过程,是一个逐步补充产品对象属性的动态推理搜索过程。这一过程是一个具有时序特征的状态转换过程。一个状态结点包括基于模型的对象描述、知识集合、设计中的基变量集合、复合变量集合、约束集合和状态评价函数等六个部分,各部分间的关系如图 5.14 所示。在一个具体的设计问题中,为简单起见,可允许基变量之间不是相互独立的,另一方面不一定去考虑所有基变量的取值。为了得到有效的设计结果,某些基变量是必须加以考虑给予赋值的,将这些基变量称为设计中的关键基变量。这样基变量集合就被分为两部分:一部分是关键基变量的集合;另一部分是非关键基变量集合。

图 5.14 设计中的一个状态结点

基于状态结点转换的多策略设计算法流程如下:

(1) 构造初始结点,即根据问题空间模型,构造关键基变量集合;根据设计者的初始约束要求,构造复合变量集合及约束集合;按一定的策略对复合变量及关键基变量排序;初始化对象描述。将初始结点放入 Open 表中。

(2) 若 Open 表为空,则搜索失败,结束。否则,从 Open 表中选取评价值最高的结点 Expand-Node 进行扩展。

(3) 按照先复合变量后关键基变量的顺序,取出 Expand-Node 变量表中第一个变量 v ,分别取其允许值(检查基变量 Nogood 表以及复合变量 Nogood 表,避免选取已发现为矛盾的值)。若其值能够在问题的对象模型中加以描述,则修改对象的状态,生成新结点;否则,将该变量的这一取值放入约束集,生成新结点。这样,根据该变量 v 为单值还是多值,可得到 Expand-Node 的一个或多个子结点 $N_i(1 \leq i \leq m)$ 。

(4) 分别对这些子结点 N_i 进行评价。

若 v 为复合变量,则首先检查其取值与其它复合变量是否不冲突。若冲突,则在复合变量 Nogood 表中记录之(若出现矛盾的复合变量子集,则以后可利用这一矛盾,直接过滤掉含有这一子集的所有状态结点),删掉该结点;若 $i+1 > m$,转(2);否则,换成下一个结点 N_{i+1} , $i = i+1$ 转(4)。

检查 v 的取值是否满足约束集中的约束。若否,则在基变量 Nogood 表中记录之,删掉该结点;若 $i+1 > m$,转(2);否则,换成下一个结点 N_{i+1} , $i = i+1$ 转(4)。

用评价函数对结点 N_i 进行评价。若该结点已不存在未赋值变量且评价值满足要求,则转(5);否则,继续。

对该结点运用知识,寻求将规则知识与该结点的对象状态描述相匹配,若激活某规则,则将推理结果视为新的复合变量,放入该结点的复合变量集合(插入适当位置)。

将该结点放入 Open 表。若 $i + 1 > m$, 转(2); 否则, 换成下一个结点 N_{i+1} , $i = i + 1$ 转(4)。

(5) 成功结束。

在这一算法的处理过程中,变量及约束均为可变的,这是由设计知识推理的动态特性所决定的。在该算法中需要注意以下两个细节的处理:一是知识推理中应避免重复匹配所造成的死循环,具体策略是在每一状态结点中,对于知识库中无变元的规则,一旦被匹配后,则在其后继结点中统一删除该规则。对于含有变元的规则,则应记录其匹配的历史情况。二是在运用 Nogood 策略时需要对变量及其约束统一编号。由于在不同结点其变量和约束可能是不同的,随着状态结点的迁移,通过运用知识,将会出现新的变量及约束。对其进行统一编号,使得相同的变量及约束具有同样的标识,才能有效地运用 Nogood 策略。注意除了初始状态时构造的基变量、复合变量以及约束外,在后续求解过程中新增的复合变量及约束均是由规则匹配所引起的,因此可根据规则号以及规则中的变元匹配情况顺序给出复合变量或约束的编号,即由相同的规则与相同的变元匹配而导出的复合变量或约束令其具有相同的变量或约束序号。为此需要根据规则号以及变元匹配情况制定一个复合变量编号及约束编号表进行编号管理。

该算法可用于机械结构优化、化工设备布局等问题,如化工设备布局包括各种塔、泵、贮罐等等,它们通常布置在露天或车间内。各种设备间存在某些位置关系约束。此外,设备之间存在着交互作用,这些作用主要是通过管道的连接来实现的。由于管道价格昂贵,因此,应考虑管道连接的费用。此外,还应考虑由操作、维修人员在设备间走动等而造成的费用。在布局中,除了应使设计结果满足各种设计约束外,还应尽可能缩小由于设备间的相互距离而造成的交互作用总费用。研究化工设备布局设计自动化,使设计结果满足一定约束,降低由管道连接及其它交互作用而发生的费用,具有重要意义。

5 3 5 基于不同粒度布局中的模糊建模及其应用

1. 不同粒度的布局系统

设计求解是一种智能行为,知识的运用是设计的基本特征之一。运用知识是缩小搜索空间、降低问题复杂度的根本途径。布局知识的表达与运用影响着布局设计自动化或智能布局辅助工具的研究进程。在模型基布局设计中,空间模型是人们从某一侧面(facet)或者某一粒度(granularity)去观察布局空间关系,所得到的抽象结构。由于角度或粒度不同,一定的模型结构一般只能表达有限的空间布局知识。因此,采用单一模型的布局系统,不可能充分描述设计逐步求精过程各阶段的设计知识。在缺乏设计知识的情况下,搜索呈现盲目性,导致组合爆炸或生成无效解。

人类智能的一个公认特点,就是能从不同粒度及不同的侧面上运用与该粒度或侧面相关的知识去观察与分析同一问题,并能不同粒度或侧面间往返自如。人类在求解复杂问题时,往往先从较抽象的层次(即较粗的粒度世界)出发,逐渐深入到其中的细节(较

细的粒度世界),这种处理不同粒度世界的能力,是人类问题求解的强有力表现。因为设计过程是一个由粗到细、不断求精、从问题的不同侧面逼近设计结果的渐进过程,所以在设计智能化的研究中应考虑如何将各个粒度侧面对问题的描述集成在一起,并提供它们之间的交互转换方法。通过各粒度层间的交互协作,求得问题的某一或某些可行解,在某些优化准则下,进一步求得问题的较优解。

多数计算机布局系统局限在某一粒度层上建立单一模型对问题求解,难以模拟人类在不同粒度上进行求解的过程。建立一种基于不同粒度的布局模型,它以多个单一粒度布局模型作为其有机组成部分,将是解决问题的一种途径。现有的一些布局模型,如墙模型、相邻图以及所研究的三维正交方位图模型、约束图模型等,分别反映了布局空间关系的某一粒度或侧面,为进一步构造这种包含多模型的不同粒度布局模型奠定了基础。

本节首先介绍研究了问题求解的粒度理论在布局建模及设计求解中的应用。针对实际问题的需要,采用了粒度投影规则方法,在粒度世界的结构非拓扑且非半序的情况下,由投影规则可建立不同粒度世界间的联系,使不同粒度世界上的结构及信息能够有效转换。以相邻图、规范约束图以及层次约束图模型为不同粒度,综合这些模型为其有机组合部分的基于不同粒度的布局模型,研究了该模型各粒度组成部分间的投影与合成关系。采取基于不同粒度布局模型的设计求解算法,构造基于不同粒度的布局系统框架。其次,本节还对模糊技术在布局设计中的应用进行了介绍。问题求解总是在一定的设计约束限制下进行的,在实际设计中,这些约束往往具有一定程度的伸缩性或某种不确定性,也即存在模糊性。这种具有模糊性的约束,称为弹性约束。本节对布局设计中的弹性约束进行了初步介绍,采用模糊约束集的一种构造方法。通过对布局中目标函数的分析,介绍了布局设计后期基于层次约束图的凸规划,以建立布局问题的模糊线性规划模型。

2. 基于不同粒度的问题描述

问题求解的现有描述方法,如状态空间法 (state-space representation) 及问题归约法 (problem-reduction representation),虽然在一定程度上反映了人类问题求解的过程,但却忽略了某些重要特点。基于不同粒度的求解方法可以弥补其不足,人类可从不同的粒度世界去分析、观察同一问题,或采用不同的粒度世界去解决不同的问题。采用粒度世界描述一个问题时,涉及若干要素,以下给出这些有关要素。

可采用四元组 $M_i = (X, \text{ , } f, C)$ 来描述某一问题的某一粒度空间。其中 X 表示问题的论域,在某工厂设备布局问题的例子中,工厂设备的全体就是该问题的论域; 是论域的结构,指论域中各元素的相互关系,如设备位置间的关系、设备之间单位距离的物流费用等; f 表示论域的属性,用函数 $f: X \rightarrow Y$ 表示,可表示设备自身的属性如重量、振动强度、噪音强度、设备类别、安放位置等,其中 Y 可以是 n 维空间的 R^n 中的集合,也可以是更一般的空间; C 表示对 $(X, \text{ , } f)$ 的约束,用于说明对设备自身的属性、设备间的关系所作的某种限定。

除了问题的粒度模型 $\{ M_i \}$ 外,还应定义模型间投影规则的集合 P 及问题的优化准则 F 。 P 为将该粒度的问题模型投影到其一个或几个商空间 (quotient space) 的投影规则,包括对论域 X , 结构 , 属性集合 f , 约束集合 C 的投影规则; F 为优化准则,可表示为函

数 $F:(X, \cdot, f) \rightarrow Y$, Y 通常为一维正实数空间 R^+ 。因此,一般问题求解可描述为 $(\{M_i\}, P, F)$, 即求结构 $\{M_i\}$, 在投影规则下, 使得诸 $M_i = (X_i, \cdot_i, f_i, G_i)$ 一致, 且使得 F 最优或较优。

采用粒度思想要借助商空间概念, 所谓商空间是由原粒度空间上的某一等价关系诱导的等价类所构成的新粒度空间。论域的任意一个划分(分类)就对应一个等价关系 R , 不同的等价关系对应问题的不同粒度, 如在以上的工厂设备布局问题中, 若把设备按其隶属的车间分类, 那么由这一等价类可诱导出以车间为基本元素的新论域上的粒度空间。在这一空间中, 同一车间的若干设备就无法区分了, 因此, 由等价关系诱导的商空间是更抽象的粗粒度空间。

3. 基于不同粒度的布局模型

(1) 模型构造 设计过程是从概念设计到详细设计不断求精的过程, 单一模型难以适应这种渐进过程中知识表示与搜索求解的要求。粒度粗的模型难以表达详细设计阶段进一步求精时的知识, 使得设计无法进一步深入; 粒度细的模型难以描述概念设计阶段的抽象约束, 搜索空间规模庞大, 易陷于搜索的组合爆炸。

为有效利用布局问题中的多个侧面所提供的约束知识, 本节在建筑房间布局问题的研究中采用了多粒度模型, 构造了一种基于不同粒度的布局模型 (granularity-based layout system, GBLM)。该模型由三个子模型构成, 分别为相邻图、规范约束图与层次约束图。实际上如果把布局结果图(即布局坐标图)也视为布局模型, 那么它是粒度最细的模型, 因此其它三个子模型均可作为它的基于某一划分的商集。该模型的基本构造如图 5.15 所示。

图 5.15 层次布局系统框架结构

在 GBLM 模型中, 根据各子模型的抽象程度不同, 将其分为三个粒度层: 第一粒度层为相邻图, 这是一种表示布置物间相邻关系的无向图模型, 图中结点表示布置物, 两个结点间的无向边表示相应布置物间的相邻关系。在不同问题中, 允许对相邻语义作不同的解释。在房间内家具布置时, 相邻可能意味着两物体之间无其它障碍物, 且其距离不超过某一值。在建筑房间布局中, 所谓相邻往往包括两房间有门相连接的关系和仅仅要求两个房间在空间上相邻, 不一定要要求有门连接两层含义。因此, 针对房间布局问题, 这里将相邻图稍加改进, 用两种无向边分别表示空间连接关系与空间相邻关系。从某种意义上讲, 相邻关系是一种比方位关系更抽象的关系, 具有较粗的粒度层次。第二粒度层为规范约束图模型。第三粒度层为层次粒度层模型。

在基于不同粒度的布局模型中,利用各个子模型不同粒度,可使设计过程逐步从粗粒度子模型空间过渡到细粒度子模型空间。在搜索求解时,首先搜索粗粒度子模型空间,即“相邻图”空间,根据该粒度空间上的约束知识及早删除搜索分枝,缩小搜索空间;在细粒度子模型空间,可以表达详细设计阶段的设计约束,使设计结果能够进一步细化。在粗粒度空间搜索中所得的抽象结果制约着对细粒度空间进一步搜索的范围与方向。

在相邻图与规范约束图中,布置物是作为一个整体看待的,因此布置物整体是其论域中的元素,可将论域中的元素逐一编号为 $1, 2, \dots, n$ 。而对于任意一个布置物 i ,在平面布局的情况下,在层次约束图中每一布置物被分解为相互有结构关系的两对边(四条边),将其表示为 $i(1), i(2), i(3), i(4)$ 。

(2) 约束知识的不同粒度表示 布局中约束知识的表示,可用一阶谓词作为基础,将谓词与布局各子模型结构间的相应影射关系(由此决定了结构、约束及推理规则的投影规则)存储在数据库中,可以在布局求解中利用该数据库所建立的影射关系,实现各子模型间的交互。在不同粒度模型中的谓词表示也不一样,如“甲物体与乙物体相邻”及“甲物体在乙物体的左方”,在不同粒度模型中的谓词表示,可通过数据库字段间的关联建立同一空间关系在不同粒度层间的投影规则。用户可根据自己的实际需要建立自己的谓词,并将这些谓词在不同粒度模型中加以表示,从而建立各子模型结构间的影射关系,这使系统具有灵活性和可扩充性。

(3) 实例分析 以下以一个住宅布局实例来说明基于不同粒度的布局算法过程。

住宅布局的空间单元数目为 9,分别为 Living Room(LR), Kitchen(K), Master Bedroom(MB), Children Bedroom(CB), WC, Bathroom(B), UB(Balcony), C(Corridor), E(Entrance)。用于描述每一空间单元的长宽及面积的最大、最小限制约束(表 5.1)。用于描述任意两个空间单元是否允许有门相互连通,可采用可能性相邻连通矩阵。若矩阵的某一元素为 1,则该元素所在行列对应的空间单元之间允许有门相连通;否则,它们之间不允许有门相通。该矩阵如表 5.2 所示。

表 5.1 各空间单元的尺寸限制

SpaceName	min area	max area	min length	max length	min width	max width
LR	14.0	20.2	4.2	4.8	4.0	4.0
MB	11.0	15.0	3.0	4.0	3.0	4.0
CB	10.0	13.0	2.8	3.8	2.8	3.0
WC	1.0	2.5	0.8	1.0	1.2	1.8
B	3.0	4.0	1.0	2.0	1.5	3.0
UB	2.0	4.0	1.5	3.0	1.0	1.0
K	7.0	11.0	2.0	4.0	2.0	4.0
C	2.0	4.0	2.0	4.0	1.0	2.0
E	1.5	2.0	1.5	2.0	1.0	1.0

表 5 2 可能性相邻连通矩阵

	LR	K	MB	CB	WC	B	C	UB	E
LR		1	1	1	1	1	1		1
K	1						1	1	1
MB							1		1
CB							1		1
WC							1	1	1
B							1	1	1
C	1	1	1	1	1	1			1
UB		1			1	1			
E	1	1	1	1	1	1	1		

在住宅布局中,对相邻连通关系存在若干内在限制,譬如,厕所的门数目不能超过一个,等等。这些限制可以用连通度及有关规则的形式表示出来。如果用 $R(i)$ 表示单元 i 与其它单元连接门的数目, $c(i,j) = 1$ 表示 i 与 j 有连接门, N_{\max} 表示某单元所允许门的最大数目,则这种限制可用表 5 3 加以表示。

表 5 3 空间单元相邻连通关系限制

I	房间	$R(i)$	规 则
1	LR	6 $R(1)$ 1	if $R(1) > 1$, then $c(1,6) = 0$
2	K	2 $R(2)$ 1	
3	MB	$R(3) = 1$	
4	CB	$R(4) = 1$	
5	WC	$R(5) = 1$	
6	B	2 $R(6)$ 1	if $R(6) = 2$, then $c(6,10) = 1$
7	C	6 $R(7)$ 3	if $c(7,2) = 1$, then $c(7,1) = 1$ or $c(9,7) = 1$ and $c(7,1) = 1$
8	UB	N_{\max} $R(8)$ 1	
9	E	N_{\max} $R(9)$ 1	if $c(9,2) = 1$, then $c(9,1) = 1$ or $c(9,7) = 1$ and $c(9,1) = 1$

其它约束用于描述空间单元相互间的位置关系约束,如表 5 4 所示。同样,其规则和常识可定为:

- if ToSouth(MB) then ToNorth(CB);
- if Adjacency(LR, WC) then Adjacency(LR, B)。
- if direct(x, y, West) then direct(y, x, East);
- if direct(x, y, East) then direct(y, x, West);
- if connect(x, y) then adjacency(x, y);等等。

表 5 4 空间单元相互间的位置关系约束

约 束 语 义	谓 词 表 达
MB 在阳面	ToSouth(MB)
B 与 WC 共用一墙	ShareCommanWall(B,WC)
MB 在 K 的东南方向	Direct(MB,K, South) Direct(MB, K, East)
MB 与 C 南北方向相邻,有门连通	HasArcC(MB (3), C (1), 0) HasArcC (C (1), MB (3), 0) HasArcC(C (4), MB (2), 1) HasArcC (MB (4), C (2), 1)

在相邻图中, 结点代表布局中的空间单元, 连接边分为两种, 一种用于表示 connect (x,y); 另一种用于表示 adjacency(x,y)。

5 4 智能 CAD 系统实例

智能 CAD 系统在机械、建筑、电子、化工设备、工艺美术等领域应用的研究, 已经走出理论研究, 投入了实际应用。以下介绍几个 ICAD 应用系统。

5 4 1 插齿刀智能辅助设计系统

高精度插齿刀设计涉及到很多复杂曲面和曲线问题, 传统的刀具设计一般要经过资料检索、方案构思、计算分析、设计绘图等步骤, 并且还要在刀具制造和使用过程中多次修改完善, 整个过程需要较多的工作量和较长的周期, 设计质量也因人和物质条件而异。由于受到设计计算能力的限制, 几乎普遍采用比较粗糙的近似方法, 影响了插齿刀的设计精度。太原工业大学机械系设计的插齿刀智能 CAD 系统, 在准确的插齿刀侧刃螺旋面数学模型基础上建立了比较完整的设计数据库, 广泛考虑到了齿形部分、结构设计部分、磨齿机砂轮修形曲线样板等设计参数的优化, 将插齿刀设计提高到一个新的水平。插齿刀设计过程既包含有大量的数值计算性问题, 还包含有方案的拟定、结构设计、材料选择、技术要求等一些创造性活动, 需要集中丰富的知识和经验, 这些工作是常规 CAD 方法难以实现的。因此, 专家系统用于普通插齿刀和高精度直齿插齿刀、斜齿插齿刀、硬质合金插齿刀等设计中, 对提高我国插齿刀设计水平有积极的作用。

1. 知识获取

一个智能系统的构造, 首先必须考虑知识的表示及获取。根据问题性质与规模, 该专家系统的知识采用谓词逻辑、产生式等表示法。在设计型专家系统中, 为完成各项子任务需设计多个子系统, 而子系统一般是采用基于规则的演绎系统, 即产生式系统。产生式系统表示大体分为事实库、规则库和控制部分(又称解释器), 其主要的控制策略是推理。本系统的推理方式采用系统先提出一个假设的反向推理方式, 并逐个测试各个子目标, 若子目标全能匹配, 原假设成立; 否则, 系统又提出一个新假设, 按同样方式进行, 直到假设为真或者再不能提出新的假设为止。

刀具设计中处理的信息有两类: 一类是专家启发性的知识; 另一类是数据和公式等规

范化的知识。因此本系统的知识库包括规则库、数据库和框架结构。

产生式规则具有如下形式：

IF(前提 1 (前提 2 ... 前提 n))
THEN(结论 1 (结论 2 ... 结论 n))

知识库中规则被控制系统调用后,便对事实库中的信息进行操作,若条件(IF)中的前提全部被现有的信息满足,则此规则被激活,执行(THEN)的全部动作。

2 . 知识库

知识库的知识一般来源于书籍、期刊、设计资料以及本行专家的实践经验。目前,知识库中的主要知识取自公开发表以及收集到的文献资料。具体作法是:将已有知识按重要程度和作用进行分类,通过逻辑演绎分出前提与结论,最后按 Prolog 子句表达格式存入知识库。知识库管理系统共有五个模块,分别实现对知识库知识的装入、存盘、浏览、删除和修改等。

系统的知识库中共有事实和规则 100 多条,根据实际应用的需要可随时添加或修改。系统的知识库必须包括关于以下全部知识,在设计插齿刀时,系统必须能自动完成以下工作:

- 交互输入初始条件:处理设计要求子空间的参数;
- 合理选择刀具的设计参数:处理设计子空间的设计变量;
- 完成设计计算:处理导出设计子空间的设计变量;
- 评价设计结果:如外啮合插齿刀齿顶变尖限制,齿轮过渡曲线、根切、顶切等满足用户的设计要求,高精度插齿刀还要求齿形造型误差等参数满足加工齿轮精度要求;
- 进行工艺参数设计计算。

该系统用产生式规则表示启发性知识,用数据文件表示各种标准数据及国标等知识,用公式表示规范性知识。

3 . 系统结构

高精度插齿刀设计的智能化 CAD 系统的总体结构如图 5 .16 所示,整个系统采用模块化程序结构的设计方法,其主要功能为:

- (1) 知识库模块 此模块收集汇总了插齿刀设计领域中的各项经验知识,包括方案论证、材料选择、结构设计等,以及由知识库开发系统建立的知识库。
- (2) 推理策略模块 根据层次知识库结构和机械设计表示方法的特点,一般来说,单独使用某一种推理方式是不能满足要求的,为此,该系统采用正向推理与反向推理相结合的方式。
- (3) 方案设计模块 此模块的作用是根据用户提供的事实和设计要求,完成插齿刀的造型,即从形式上选择采用盘形插齿刀、碗形插齿刀、锥柄插齿刀,从齿轮形式上选择直齿、斜齿,从工件方面选择内齿轮插齿刀或外齿轮插齿刀,从使用方面选择标准插齿刀、专用插齿刀、剃前插齿刀,从齿轮硬度方面选择软齿面、硬齿面等插齿刀等。同时还要根据要求,决定插齿刀参数的设计原则,这基本上是一个推理过程,由推理机和知识库配合完成。

图 5 .16 插齿刀设计 ICAD 系统的总体结构

(4) 方案评价决策模块 根据程序运行的结果和设计的目标,按照插齿刀的设计指标进行综合评价,以决定方案的取舍。如外啮合插齿刀设计时齿顶变尖的限制,被切齿轮过渡曲线干涉、根切、顶切的限制,齿轮径向间隙等限制。高精度插齿刀还要求齿形误差满足加工齿轮精度要求,对大模数齿轮插齿刀,如果齿形优化后造型误差大于齿轮公差的 $1/2$,则需对刃磨齿形的磨齿机砂轮廓形进行优化设计,以满足加工要求。

(5) 方案分析和修改模块 该模块是在方案设计评价、决策不能通过时,分析失败的原因,指出修改方向和措施。根据分析失败的信息和利用知识库的有关规则进行重新设计。

(6) 输入/输出(I/O)模块 该系统利用了 TurboProlog 与 Fortran 语言的接口,其主控程序仍为 Prolog 程序。这种利用 Prolog 程序进行自动推理和 Fortran 程序进行高速计算,可以解决目前人工智能领域中不断出现的大量计算问题。

(7) 知识库开发系统 从图 5 .12 中可以看出,知识库开发系统与推理策略模块是与主程序分离的,这有利于系统的研究开发、扩展和维护,也有利于应用。但还将加强研究知识库知识的完整性、一致性和自学功能,以及与 CAD 方法的接口问题,使之成为一个比较适用的机械设计专家系统开发的工具系统。

4 . 系统功能和特点

该系统除具有普通插齿刀设计专家系统的功能外,主要是在系统应用软件中增加了有关高精度插齿刀设计齿形部分的数值计算程序、齿形优化设计程序、修形曲线样板加工参数计算和微机控制磨齿机砂轮修形系统输入参数计算程序等,可在国内现有工艺技术装备水平上,提高插齿刀的齿形精度。

插齿刀设计 ICAD 系统是一种层次化、多模块化结构,因此有利于系统的调试、补充和修改。该系统不仅用 Prolog 语言建立相应的规则库和推理机,层次之间和功能模块之间的链接由专门的接口完成,而且还具有插齿刀设计的有关经验图表、标准数据、规律总

结等的存储能力。Fortran4.0 完成数值计算,并用接口实现两种语言程序的相互调用。绘图由 Fortran 和 AutoCAD 完成。系统能自动完成普通插齿刀和高精度插齿刀的设计、计算、编写说明书和绘图等工作。整个系统的用户接口友善,可实现人机对话,在运行期间还可对数据进行增添、删除、修改等操作。

该系统利用了 Prolog 的窗口及菜单技术,是一个具有智能化的用户接口,用户界面较为友好。该系统是一个多窗口的功能菜单选择,可进行如下操作:进行设计选择,调用帮助信息,执行知识库管理操作,信息窗口随时显示有关操作、计算执行步骤等活动。对话窗口在系统运行时接受用户的输入并显示有关计算结果、推理结论。

5.4.2 铣刀设计的智能 CAD 系统

目前在机械行业中已出现了不少关于复杂刀具的 CAD 系统和一些不太完善的专家系统,但主要都是针对齿轮刀具的,而对于理论尚未完全定型的铣刀、拉刀等刀具,这方面的工作却开展得很少。以下介绍的四川机械设计院的铣刀设计专家系统,它能完成普通铣刀设计的全部过程。

1. 普通铣刀智能 CAD 系统的组成

铣刀设计智能 CAD 系统分为控制和资源两大部分。控制部分是放置刀具设计所共有的东西的,被设计成模块化形式,这样只需修改很少的地方就能找出控制新刀具的设计。控制部分包括初始参数的输入模块、方案选择模块、方案的解释模块。资源部分包括知识库、数据库、初始参数数据库、解释知识库等。控制部分与资源部分完全分离开,所以它可以在互不影响的情况下,修改系统的控制结构,或增加新规则和更新知识库,使系统更趋完善。

由于初始参数是每种刀具设计前都必须首先输入的,为了使其具有良好的通用性,需将其编制成固定模块。在设计另外一种刀具时,只对初始参数数据库和检测知识库进行适当的补充和修改即可。在初始参数的输入模块中,只需按提示来输入数据,当参数输入完毕后,系统可模拟人的推理思维,对这些参数进行自动检测,如发现有超界或错误,例如粗糙度值或精度等级超界等,则会自动将这些参数提示出来,请求用户修改。

在铣刀方案知识库中,收录了大量确定铣刀方案的知识。系统根据已知条件在屏幕上显示一种或几种方案。几种方案的排列是以其可信度值的大小而定,用户可以从这些结论中选择可信度最大的一个作为推理结果,也可以根据实际生产情况,从中选取符合加工要求的结论。

经过初始参数的输入,从系统的推理而得到所应选用的铣刀类型后,就可以调用铣刀图形库,利用参数设计法,绘制出理想的铣刀工作图纸。因此该系统具有较为友善的人机界面,设计中,用户可随时输入或检查设计参数,并且系统能对其推理作出详尽的解释。

2. 知识的表示与组织

知识的表示与组织是建造专家系统的基础,直接关系到系统的推理能否模拟出专家的思维过程。在铣刀设计中,所涉及到的知识分为经验规则型和计算公式型两类,并且采

用产生式规则描述设计经验、判断决策方面的知识;用框架结构表示数据类知识和推理过程中的一些知识记录,并力图建立设计的规范语言来支持这些知识描述。

(1) 铣刀设计知识语言集 某一领域里的知识总可以用一些规范的语言单位,按照一定的语义去组成各种形式的知识。实际上设计过程中的事实、概念和术语就是一些基本的语言单位,它们一方面表示了一定的知识,另一方面又可用它去构造其它知识形式。该系统以 AutoLISP 语言中的表为基础来描述它们,分为关联组元表示、关系谓词表示、函数谓词表示等三类。

关联组元表示方式。由于在铣刀设计中一些孤立事实,术语常用实体、属性和值的概念以关联四元组的形式表示,其形式是:

(Object Attribute Value Unit)

这里的实体 Object 是客观存在并可以区别的事物,属性 Attribute 是实际具有的某些特征,值 Value 是一个广义的概念,可以是数值、字符串或表。如果是数值,有时需带上相应的单位 Unit。在铣刀设计中,为了便于运算,实体有时用代码符号表示,结果输出时,往往又需要对此进行解释。这种解释是用二元组构成的设计信息字典集来实现的,例如:(Width“宽度”), (Length“长度”)等,就是用 Width 代表被加工零件的宽度,Length 代表零件的长度。

谓词是表示思维对象的性质或多个对象之间的关系的词,在表示事实等知识时,常用谓词演算中的关系谓词来表示,型式为:

(RP Object Attribute Value)

三元组。其中 RP 是表示对象与属性或值之间关系的谓词, Attribute | Value 表示或的关系,例如:(KNOWN Width Length)表示变量 Width 和 Length 是已知的。上述表达式的特點是以谓词为中心,对象属性或值是可变的。

上述两种表示形式用于描述事实的状态,但还不能完全描述控制和操作方面的知识,因为有些子句描述的事实不是静态的,需要根据谓词的意义进行一些解释和动作,以实现信息的检索、匹配、修改和存储。这类子句称为函数谓词表示,其形式为:(FP Parameters),式中参量 Parameters 的含义较广,可以是 LISP 内部函数组成的表达式或谓词表示的关联组。FP 为规定某一函数谓词所要执行的特定操作。

(ORL 条件 1 条件 2 ... 条件 n)

(ORF 条件 1 条件 2 ... 条件 n)

上述两式均用于从左到右判断 条件 i 是否成立,但操作方式却有差别。第一式用于判断由 LISP 内部函数写成的逻辑表达式是否成立,若成立则返回 T , 否则返回 NIL ; 第二式的条件是符号表示的事实,需要到综合数据库中去查询和匹配,条件成立返回 T , 否则返回 NIL 。之所以这样划分,是因为设计中事实处理属于符号匹配,数值计算和条件不等式属于求值。可见函数谓词的作用是告诉推理机按约定的意义,以不同的方式完成操作,这种处理方式也是为了适应设计知识集内符号推理和数值计算交织进行,有时甚至是为了在同一条规则中出现的需要。

将上述三种形式归结为铣刀设计知识语言集的内容,它们都是设计中常用的术语、事实和概念的简写形式。实际上,它不仅适用于铣刀设计,其具体内容还可以根据设计对象

的不同而不同,它们都可以作为描述其它各种知识的基元。

(2) 知识的产生式规则描述 根据铣刀设计特点,铣刀设计知识采用产生式规则描述,并对其进行分类处理,以兼顾事实推理与公式计算的需要。

事实规则是,设规则的前提 P 和结论 A 均是事实描述的集合,记为 $P \rightarrow A$,它用来表示设计原则、经验和一些启发性知识片断。其特点是前提 P 与综合数据库 FACTS 表匹配成功后,按一定的冲突解决策略行动,并把结论置于 FACTS 表中。

计算规则是设参数集合

$$X = (x_1 \cdot x_2 \cdot \dots \cdot x_n),$$

计算规则表示为

$$(V_{x_i}) [P \rightarrow X] \rightarrow F = f(x)。$$

此式含义是前提事实集合 P 成立,并且参数集合 X 中的所有元素已知时,则按 $F = f(x)$ 的关系计算参数,这类规则主要用于描述设计中需要用的一些计算,以帮助决策的情况,计算公式都比较简单。

判断规则表示某一规则的前提事实集合和参数集合已知,并且参数集合中每一元素满足一定的约束条件时,由此规则可导出结论 A 或按 $F = f(x)$ 的关系进行计算。这类规则在设计型专家系统中经常涉及。

赋值规则的表达式为 $P \rightarrow (A = C)$,含义是某一规则的前提事实 P 成立,那么就给变量赋常值 C 。

在设计铣刀过程中,上述几类形式都有可能交替出现。在推理过程中利用这些规则,其中事实规则是一个符号匹配问题,后几类规则不仅有符号匹配,而且还要检验计算中涉及的参数是否已赋值,或不等式判断,并把计算结果传递到后继规则中去。

3. 系统的推理机

在铣刀设计中,大多是根据用户提出的条件和市场信息,设计人员利用其经验知识,逐步构造出设计方案并绘制设计图纸,是个正向工作过程。因此采用基于综合数据库的控制结构,以正向推理为推理策略,并且能进行精确推理和不精确推理。其整个推理过程为匹配、冲突解决和操作。

在推理过程中,任何一条知识的应用几乎都要同综合数据库发生联系。因为提取或存贮信息,设计的阶段结果要传递到后续过程作为初始条件,所以综合数据库是一个全局可访问的工作空间,用以记录问题的初始数据和推理过程中间结果及最后结论。它的组织是否合理,对提高推理效率,顺利进行设计具有直接关系。由于本系统采用功能较弱的 AutoLISP 语言,给设计综合数据库带来了一定的困难。

在铣刀设计智能系统中采用了两个线性表和一个框架来组织动态信息。动态事实表 FACTS 记录初始事实中的事实条件和执行规则获得的事实结论;动态数值结果记录框架 Results 主要用于记录初始事实中的非事实条件和执行规则获得非事实结论。在本系统中,对于被加工零件的外形描述就被处理为事实条件,而加工零件的精度要求、表面粗糙度要求,则被表示为数据条件。之所以这样分类处理是为了进行匹配,而计算规则和判断规则等则只需在 Results 表中进行就可以了,不需对整个综合数据进行扫描,从而使匹

配操作被限制在某一个局部的空间中进行,提高了效率。用于记录每次推理成功应用的规则编号表具有两个作用:一是避免推理机反复调用已成功应用的规则;二是把设计过程的记录,用于今后实现解释、咨询等功能。

4. 系统的深层解释

解释能力是专家系统具有的重要特性。刀具设计是一门涉及面广、经验性强的学科,它没有太多的现成法规,许多知识都是设计人员在长期的生产实践中摸索出来的。因此完善的解释机制,对于提高系统的可接受性及设计知识的继承和发展尤为重要。

目前常用的方法有文本预置法和追踪解释法,这些方法尽管能对系统行为和当前动作的目的作出解释,然而却不能对系统行为的合理性作出解释。于是该系统根据专家潜意识中隐含的智能行为,提出了一种新的解释方法——深层解释。例如在普通铣刀智能 CAD 系统中有这样一条规则:“如果加工平面经过淬硬、表面粗糙度 $Ra \leq 1.6$ 、精度 $IT \leq 7$,则采用硬质合金端铣刀,其可信度为 0.95”。如果用户对此规则的结论提出疑问,而系统仅将此规则的内容向用户解释,那显然是不够的。也许用户对这条规则本身的合理性就存在怀疑,他可能提出这样的问题:“为什么采用硬质合金端铣刀而不采用高速钢端铣刀或圆柱铣刀来加工该平面?”对于这一类问题,用追踪解释法是不能作出任何解释的。然而,专家在使用这条规则时,在他的潜意识中却存在某些理由说明这条规则的合理性,对于上述问题,专家可能作出如下的回答:“虽然用高速钢端铣刀或圆柱铣刀也能加工平面,但由于被加工的工件是淬硬工件,而加工表面质量要求高,与其余两种相比,硬质合金端铣刀除能满足上述要求外,还可以实现高速切削,提高加工生产率,所以应采用硬质合金端铣刀加工该平面。”

正是由于专家潜意识中的这些理由,才使人们相信专家所作的结论,这也是使得这条规则进入专家系统知识库的依据。尽管这些理由很难收集,有时专家自己也说不清楚,然而一旦收集到关于规则存在的合理性和理由,那么,我们就可以向用户提供更深一层的解释。

铣刀设计智能 CAD 系统克服了以上缺点,在其知识库中不但存储了求解问题的规则,同时也存储了大量的关于规则的解释。这些有的是专家的,有的是由于规则的规范性要求而使规则的意义不很明显时,为了说明规则的含义而附加上去的。这种深层知识的存储,不但能向用户提供深层解释,也使得规则的内部表示简单化。当用户需要对某一结论作更进一步的解释时,尽量从用户的角度出发,将每一种可能的方案(无论是否合理)都显示在屏幕上,同时最佳方案也显示在屏幕上。用户可以从诸多的方案中,选择他感到不清楚、有疑问的方案,向系统提出问题,系统将作出深入、完善的解释,从而使系统同时也具备了一定的学习功能。

机械设计实质上是在满足一定条件下的结构选取及参数计算的过程,各种零件的设计具有一定的共性。为了便于今后的移植推广,铣刀设计智能 CAD 系统的建造以专家系统外壳为努力目标,即系统本身并不嵌套有关铣刀设计的特定信息,所有设计变量及特征参数均以数据文件的形式存放。这样,只要更换知识库及变量名,就能将之应用于机械设计的其它领域。

本系统利用 AutoLISP 语言作为建立专家系统的工作语言,实现了与 AutoCAD 绘图环境的统一,使推理求解、数值计算和绘图一体化,提高了微机智能 CAD 系统的实用性。

5 4 3 机床夹具设计专家系统

机床夹具的设计是机械制造业中生产准备工作的重要部分。现代产品朝着多品种、小批量的方向发展,产品更新换代很快,有时一个新产品上马需要十几套乃至几十套夹具,但用于夹具设计和制造的生产准备时间太长,耗费人力和物力,还可能贻误市场良机。同一个夹具,不同的设计人员将设计出全然不同的结构。夹具零部件的重复利用率很低,库房里堆满了用过的旧夹具,而新夹具产品要上马时又苦于不能尽快地设计制造出来。由此可见,实现机床夹具的计算机管理,以重复利用夹具零部件的规范化设计为目标的计算机自动设计已成为急待解决的问题。

这里介绍的是新疆工学院设计的机床夹具设计专家系统,它是一个在 VP-Expert 专家系统开发工具上开发的,它可完成各种普通车、铣、钻、镗、磨、刨、插、拉、齿轮加工机床上的专用夹具的自动设计。

1 . VP-Expert 专家系统开发工具简介

夹具智能专家系统用中文之星 1.2, Windows3.1 和 VP-Expert 专家系统工具作基本开发环境,主要是利用它的内存自动管理、同时运行多个程序(多任务)的功能,使有的程序在前台运行,有的程序在后台运行,根据程序的具体情况确定其分享 CPU 的相对时间(Priority time),并开发出一些图文并茂的用户界面和实现数据的动态交换(DDE)。

VP-Expert 属于规则型专家系统开发工具。由于机床夹具设计中的知识基本都是确定型知识,完全可用 IF...THEN...的形式表示,所以使用这个专家系统开发工具很合适。这样不仅大大节省了建造专家系统的时间,而且建造的专家系统效率高、运行速度快,其知识库还易于扩充和能根据不同的用户(工厂)灵活变动。

在 VP-Expert 中可用数据库和文本文件作为其知识库中的信息库。用 FoxPro 2.5 数据库构造了机床夹具设计专家系统中的信息库。信息库中存贮系统所用的各种参数,例如一个有关某个标准零件的记录,既包括所有零件的结构参数,又包括图纸标注的有关文本。同时,设立了专门的文本文件用来存储常用的工程术语和相应的屏幕提示信息。

AutoCAD 11.0 引进了图纸空间概念,可以更方便地生成各种视图,用它生成夹具的装配图和零件图以作为指导夹具制造的工程图纸。系统让 AutoCAD 主要在前台运行并使其分享大部分 CPU 时间。为了提高系统的运行速度,采用两种语言共用方式,即 Autolisp 编制程序,用 C 语言编写复杂的计算程序。

2 . 系统设计思想

(1) 典型夹具的选型 在人们试图重复利用夹具零部件的努力中,有两种解决方法:一是利用成组夹具,二是采用组合夹具。成组夹具是建立在成组工艺之上的,没有成

组工艺就谈不上成组夹具;组合夹具则必须先购进全套组合夹具零件,一次性投入资金太大。这两种方法均有其局限性。在建立夹具智能 CAD 系统时,把每种机床上的夹具按结构分成若干大类,每大类又通过筛选,优化组合分为若干个典型夹具模型,每个典型夹具又分为大、中、小几个规格。这些典型夹具模型汲取了成组夹具和组合夹具的优点,但它又不是成组夹具或组合夹具。它是一种一定范围内的通用可调夹具和拼装夹具。只要对其中的零部件作相应变动,它就能适应一定范围内的各种工件的加工。所有的典型夹具,已能覆盖绝大部分工作,只要工件在某一典型夹具的加工范围之内,系统即能完成该夹具的自动设计。对于形状奇异的工件,系统则按定位、夹紧原理等规则推理逐一设计出相应零部件,由操作者确认,即由推理与人机交互相结合的方式完成设计。

(2) 三维设计与绘图 整个设计过程都是在三维空间进行。从一开始输入工件外形就是建立三维模型,插入夹具零部件也是三维实体,最后生成整个夹具的三维模型。图纸的输出也不拘泥于传统的三视图,而是从不同角度观察夹具及夹具零部件,选择最佳视点,经消隐处理输出轴测图(立体图形),辅之以适当的向视图或剖视图。这样的图直观、立体感强,省去了由二维的三视图想象出三维实体的费脑筋的过程,这样的图纸乐于为工人所接受。

(3) 图形函数及其属性 国标夹具零部件、国标紧固件以及典型夹具中的零部件,事先都用 AutoLisp 或 C 编程,定义为图形函数,设计时从信息库中提取数据参数赋值,生成所需的零部件。建图形函数时考虑到将来生成装配图和零件图的需要,用条件转移语句在函数内设几个分支,如装配图上零件不标尺寸,零件图上全剖或局部剖,装配图上该零件也可能不剖等。另外,将零件图和装配图上的有关文字信息分别存在不同的层(layer)或属性(ATT)中,以供将来需要时显示。

3. 系统结构

机床夹具设计专家系统流程图如图 5.17 所示,它分为车床夹具设计、铣床夹具设计、钻床夹具设计、镗床夹具设计、磨床夹具设计、刨床夹具设计、拉床夹具设计、齿轮加工机床夹具设计、专业化机床夹具设计九个子系统。每个子系统由工件输入模块、定位误差分析计算模块、夹紧力校核模块、零部件插入模块、夹具体设计模块、装配图生成模块、零件图生成模块组成。工件码按 JLBM-1 码由菜单提示选择输入。信息库又细分为机床设备数据库、刀具数据库、夹具管理数据库、定位元件数据库、夹紧元件数据库、对刀导引元件数据库、夹具体数据库、国标紧固件数据库等。规则库根据不同模型的需要由若干文件组成。

系统设计模块的功能是在第一次进入系统时按系统配置的硬件进行设置的,以后则是按要进入哪一子系统对软件进行配置。工件外形输入模块采用人机交互方式由键盘和鼠标输入或由数字化仪输入,建立工件的三维模型。在屏幕上显示确认无误后,提取 DXF 属性,并标识出定位面、夹紧面,提取 E 属性。

定位误差分析计算和定位元件的插入同时进行,系统根据定位面选出定位元件,并进行定位误差的尺寸链计算,结果返回屏幕。夹紧元件的插入和夹紧力校核同时进行,然后再插入其它零部件和夹具体,最后生成装配总图和零件图。整个设计过程有自动方式和

图 5 .17 机床夹具专家系统流程图

半自动方式两种工作模式。所谓“自动”就是专家系统根据推理自动完成设计,“半自动”是通过选择屏幕图标菜单,回答提示等完成设计。这两种模式可在设计的任一阶段自由切换。

第 6 章 计算机辅助工艺设计及其智能化

6.1 简介

6.1.1 工艺设计

机械制造是一种离散的生产过程。机械制造的基本特点是按照设计要求和工艺要求对毛坯进行加工。

工艺设计是根据产品设计给出的信息和企业的设备、技术等具体情况进行产品加工方法和制造过程的设计。

一般来说,工艺设计包括以下内容:

- (1) 确定工艺路线;
- (2) 工序设计;
- (3) 选用机床;
- (4) 选用刀具;
- (5) 夹具选择;
- (6) 确定工艺参数(加工余量、切削用量等);
- (7) 确定刀具路径(编制 NC 程序);
- (8) 确定工时定额。

工艺设计是联系设计和制造的重要环节,其质量的好坏将直接影响产品的质量、生产率和价格。一个好的工艺设计应满足以下几点要求:

- (1) 应能保证满足产品设计提出的所有质量要求;
- (2) 应能保证较高的生产效率,能按计划交付产品;
- (3) 生产成本要尽可能的低;
- (4) 应能保证有一个好的工作环境;
- (5) 应尽量采用先进的加工技术。

在工艺设计中,需要解决的一个首要问题是如何达到设计所提出的精度要求,在制定工艺时,应尽可能避免单纯通过依靠操作者的个人技巧来达到所要求的精度。

在工艺设计时,还必需考虑技术和经济两方面的因素,如采用何种毛坯?是采用铸件还是锻件?是否有必要采用比较先进的加工方式或设备?是否需要采用特种加工手段等一系列问题,从而达到较高的生产效率并降低生产成本。

在进行工艺设计时,有时还要把各种工艺方案进行比较,以便从中找出最好的一种。当然,这里所说的最好的一种并不一定要采取最先进的加工手段或追求最高的生产率,而必须从各方面综合考虑,如质量、成本、效率等,从中找出最好的一种。

传统的工艺设计的主要信息来源是工程图纸和技术条件、零件的毛坯图、计划产量、现有设备和工具以及有关的工艺手册等。

零件图及其技术条件是工艺设计的基础,其中应该包括以下几方面信息:

- (1) 零件形状,其中应包括必要的视图、剖面图和尺寸标注;
- (2) 技术条件,其中应包括所有必要的尺寸公差、形位公差、表面粗糙度以及其它说明;
- (3) 零件的材料、热处理以及毛坯种类。

零件的毛坯图是根据零件形状,按零件的机械加工要求和成本而设计的。毛坯的种类很多,有铸件、锻件、焊接件和棒料等。在进行工艺设计时毛坯图是设计的内容之一,同时也为下一步工艺设计提供了必要的信息,如选择定位、夹紧方式和加工工序、加工余量等。

生产批量和加工方法之间有密切的联系,因此,生产批量在很大程度上决定了加工工艺。生产批量一般可分为单件、小批量和大批量三种。

工艺设计在很大程度上取决于生产设备。工艺设计有两种情况,一种是在已有设备基础上进行工艺设计,另一种是在工艺设计的同时选择全新的设备。

其它有关信息包括各种工艺手册、工装手册、车间平面布置、车间操作人员情况和车间运输装置等。

6.1.2 工艺设计标准化

由于工厂的产品种类、生产环境、人员和经验的不同,因此各个工厂在工艺设计上存在很大差异。对于大批量生产的产品,如汽车零件、轴承、发动机零件等,一般比较有完整的工艺文件,但对于中、小批量生产,一般只采用通用设备和万能工具进行加工,工艺人员只是简单排一下加工顺序,根本谈不上工艺的标准化。然而,随着数控技术、柔性加工技术和计算机集成制造技术的发展,要求采用计算机辅助工艺设计并逐步过渡到工艺设计自动化的越来越多。这样,工艺设计的标准化和优化就是必要的先决条件。工艺设计标准化的目的,就在于使同类零件能够在相同生产条件下采用经过优化的标准工艺,并且用规范的语言来描述这一过程。这不仅可以防止不必要的工艺多样化,提高加工质量和工效,而且为实现统一的 CAPP 开发工具和 CAPP 专家系统的知识库打下了良好的基础。

加工工艺规范化的内容很多,根据目前我国在这方面的研究,大致可分为以下几个方面:

- (1) 零件加工工艺术语;
- (2) 零件加工分类规范;
- (3) 工艺路线规范;
- (4) 毛坯选择规范;
- (5) 零件热处理规范;
- (6) 毛坯余量;
- (7) 加工工序间余量;
- (8) 加工类型及加工精度选用规范;
- (9) 机床规格选择规范;
- (10) 加工顺序安排规范;
- (11) 装夹方法选择规范;

(12) 工具、量具、辅具选用规范。

6.1.3 计算机辅助工艺设计(CAPP)发展概述

机械加工工艺设计过去一般由工艺人员来完成,由工艺设计人员根据零件的材料、几何形状、尺寸、公差、表面粗糙度、硬度等产品信息,结合当前工厂的生产条件和生产批量作出工艺方案并完成工艺路线、工艺装备、工艺参数和工时定额等各种工艺文件。传统的工艺设计方法要求设计人员具有丰富的生产经验,熟悉企业的设备条件和技术水平,了解各种加工规范和有关规章制度。总之,一个好的工艺设计人员必须经过多年实践工作的锻炼。

采用人工进行工艺设计,不但需要一大批有丰富工艺知识和多年实践工作锻炼的工艺人员,而且还要采用手工方式准备各种工艺文件,工艺人员的工作主要是查手册、计算数据、填写工艺卡、画工艺图等一系列手工操作。工艺设计所需的时间多,工艺规程的可改性差,工艺设计的人为因素和随机性也较大,不利于生产率的提高。计算机辅助工艺设计能在计算机的帮助下迅速编制出完整而详尽的工艺文件,工艺人员无需重复查阅各种手册和规范,不再用手工编抄各种表格,大大提高了工艺人员的工作效率。

国外早期的 CAPP 研究工作开始于 60 年代末期,较有代表意义的是 CAM-I(computer aided manufacturing-international)系统。设在美国的计算机辅助制造国际组织于 1976 年推出了 CAM-I's Automated Process Planning 系统,取其字首第一字母,称为 CAPP 系统。虽然与现在缩写语 CAPP(computer aided process planing)所代表的词语有所不同,但计算机辅助工艺设计称为 CAPP 已为大家所公认。早期的 CAPP 系统实际上是一种技术档案管理系统。它把零件的加工工艺按图号存放在计算机中,在编制新的零件工艺时,可从计算机中检索出原来存有的零件工艺,有的可以直接使用,有的必须加以修改,如果检索不到则必须另行编制后存入计算机。

派生式系统(variant system)是以成组技术为基础,把零件分类归并成族,零件族的划分可采用直接观察法、分类编码法和工艺流程分析法等方法来进行,并制定出各零件族相应的典型工艺过程。在使用这种 CAPP 系统时,首先根据零件特点按零件族的划分方式将需要编制工艺的零件进行分类,然后从计算机中调出典型工艺,再通过文件编译修改生成零件工艺。这类系统的主要缺点在于其针对性较强,使用上往往局限于某个工厂中的某些产品,系统的适应能力较差。派生式 CAPP 系统由于开发周期短、开发费用低,因而在一些工艺文件比较简单的中小工厂中比较受欢迎。

为了克服派生式 CAPP 系统的缺点,许多大学和研究机构纷纷开展了创成式 CAPP 系统的研究工作。创成式 CAPP 系统中不存储典型工艺,而采用一定逻辑算法,对输入的几何要素等信息进行处理并确定加工要素,从而自动生成工艺规程。

从 80 年代中期起,CAPP 专家系统的研究得到迅速的发展,每年在有关杂志和国内外有关学术年会上都有大量 CAPP 专家系统的论文发表。基于知识的计算机辅助工艺设计系统已成为 CAPP 的重要发展趋势。

专家系统作为一种工具,它具备某一领域的知识并利用这种知识来解决这一领域中的问题的本领。采用专家系统的办法很适合于解决工艺设计中若干方面的问题。传统上,这些问题主要有两个主要特点:其一是工艺知识主要依靠经验,这些专门经验一般要

通过较长时间的积累;其二是这些知识本身有一定的不确定性,且随企业条件的不同而变化。因此,如果利用专家系统来进行自动工艺设计,就必须和工艺师一样收集、表达和利用这些工艺知识,而专家系统本身已为这些工作提供了很好的框架。由于专家系统的知识库容易进行修改和扩充,因此比较容易适应企业的生产环境。目前,各国学者正致力于使 CAPP 专家系统达到进一步实用化和工具化。

有必要指出的一点是,以上所述的三种系统虽然采用了不同的方法,但所采用的技术则是可以互相借鉴的,如在获取专家知识时往往需要将零件进行分类,以便按其类别来收集和归纳专家知识,人工智能的方法也可用于成组编码,从而实现零件编码自动化。目前出现的一些混合式的 CAPP 系统,正是这几种方法综合运用结果。

6.1.4 派生式 CAPP

派生式 CAPP 系统是在成组技术的基础上,按零件的结构和工艺的相似性,利用分类编码系统将零件区分为若干零件加工族,并以零件族的复合零件(也称样件)为代表编制零件的成组工艺;或利用综合路线法来编制覆盖一族零件的成组工艺,作为该零件族的标准工艺,以文件的形式存储在计算机中。在编制新零件的工艺规程时,首先根据零件信息确定其成组编码,在用计算机判别出它所属的零件族以后,调出该族的标准工艺,并根据新零件的几何及工艺信息,采用人机对话的方式,删节和编辑该零件的加工工艺,同时进行切削用量和工时定额的计算,最终设计出新零件的工艺文件。

派生式 CAPP 系统的研制可按以下步骤进行:

1. 建立一个分类编码系统

派生式 CAPP 通常需要建立一个分类编码系统,因为分类编码为识别零件的相似性提供了一种方便的手段,具体的分类编码方法我们将在下一节中具体介绍。虽然目前市场上已经提供了不少分类编码方法,但仍然有不少企业宁可根据自己企业的情况自行研制分类编码。在这里,我们建议,在各企业自行研制分类编码系统以前,应对已有的分类编码方法进行仔细研究和比较,然后尽可能采用比较成熟的分类编码方法。这样做可以节约大量人力物力。如果无法找到合适的分类编码系统,建议在自行建立分类编码系统时,请有关的专业人员来进行指导。原则上,分类编码应覆盖本企业的所有产品,其中应包括企业准备发展的产品。因为有时无法预测企业要开发何种产品,故建议编制分类编码系统时要带有一定柔性,例如,留几个备用码位以便今后扩展。

2. 零件族的划分

成组技术提出了零件族的概念,把结构形状相似、尺寸相近和工艺相似的零件归并为一个加工族。

工艺相似使同一类零件可以采取相类似的工艺路线,这是制定标准工艺的前提,而结构相似则是工艺相似的基础。尺寸相近可使零件在同规格的加工设备上加工作。在三个相似中,最重要的应该是工艺相似。

零件族的划分方法主要有以下三种:

(1) 经验法 它主要根据有经验的老工艺人员多年来的工艺经验,按零件的形状、大小、精度要求并考虑零件功能特征的相似性划分零件族,例如,回转体零件可划分为齿轮类、套类、三角皮带轮、螺母类等,然后再把每一类零件细分为零件族。这种方法全凭经验,简单易行,适合产品变化不大的企业。

(2) 工艺流程分析法 工艺流程分析法是把工艺过程相似的零件划分为一个零件族,也可对用上述方法定出的零件族进行工艺流程分析,把个别工艺流程特殊的零件剔除或修改其工艺,最终确定零件族范围。

(3) 分类编码法 采用这种方法首先要选择零件分类编码系统,最好选用已有的比较成熟的系统,如我国的 JLBM-1 系统(JB/ Z251-85)、德国的奥匹兹系统和日本的 KK-3 系统等。如果已有的系统确实无法满足要求,或者本企业产品品种单一,例如轴承厂、标准件厂、液压件厂等,则可对已有的编码系统进行修改或另行编码。然后将现有的零件利用编码系统进行编码。分类编码法按零件的成组编码划分零件族。划分零件族的方法可分为以下两种:

特征位数据法。根据所采用的分类编码系统,确定有代表性的几个码位作为划分零件族的特征码位,如果一个零件的成组编码全部落在某零件族的特征码域内,表示这个零件属于该零件族。例如有三个零件,其特征码分别为:

04306 3070
04100 3072
04102 3072

如果选用第 1,2,3,4,9 码位为特征位,那么,很明显第二和第三个零件为一组。具体选用哪些位作为特征位,可根据实际情况确定。

码域法。码域法是从编码系统的各码位中,人为地选出若干个特征项作为划分零件组的依据,表 6 .1 为某个轴类零件的码域表(采用奥匹兹编码系统):

表 6.1 码 域 表

项次 \ 码位	1	2	3	4	5	6	7	8	9
0									
1									
2									
3									
4									
5									
6									
7									
8									
9									

表中各码位中划勾的项次是人为选定的。

第一码位选 1, 2 两项($0.5 < L/D < 3$ 和 $L/D > 3$);

第二码位选 0, 1, 2, 3 项(外形光滑和单向台阶);

第三码位选 0 项(无内孔);

第四码位选 0, 1, 2, 3 项(无平面加工或外部有平面加工);

第五码位选 0 项(无辅助孔);

第六码位选 0, 1, 2, 3 项($D \leq 150$);

第七码位选 2, 3, 4, 5, 6 项(钢件);

第八码位选 0, 1 两项(棒料);

第九码位选 0, 1 两项(无精度要求和外形有精度要求)。

只要代码落入码表中划勾项次的零件就可以划在同一零件组内。采用码域法的主要目的是为了扩大成组的批量,在码域表中选带勾的项次的主要根据是从加工相似的角度来进行考虑的。如第一码位中选用了 1, 2 两项而没选用 0 项,这是因为 0 项的含义是 $L/D < 0.5$,属于通常所说的盘类零件,从加工角度来看,显然与第 1, 2 两项的轴类零件区别较大,放在一个零件族内不合适。由于各企业的条件不同,因此很难规定一个通用的选项标准,各单位需自行摸索和选定,有时还要反复几次,才能将码域表搞得理想和适用。

3. 编制典型成组工艺

(1) 复合零件法 复合零件法又称样件法,它是利用一种所谓的复合零件来设计成组工艺的方法。复合零件可以选用一组零件中实际存在的某个具体零件,也可以是一个实际上并不存在而纯属人为虚拟的假想零件。复合零件应是拥有同组零件和全部待加工表面要素或特征的零件。按复合零件设计成组工艺,只要从成组工艺中删除某一零件所不用的工序(或工步)内容,便形成该零件的加工工艺,见图 6.1。

实践证明,对于形状较简单的回转体零件,用复合零件法编制工艺比较合适。

(2) 复合路线法 有些结构比较复杂的回转体零件组和绝大部分非回转体零件组,因其形状极不规则,所以要虚拟复合零件相当困难,这时可采用复合路线法或称流程分析法。复合路线法是在零件分类成组上,把同组零件的工艺文件收集在一起,然后从中选出组内最复杂,也即最长的工艺路线作为加工该组零件的基本路线。

复合零件法比较适合结构比较复杂的回转体零件和大部分非回转体零件。

6.1.5 创成式 CAPP

创成式 CAPP 与检索式和派生式 CAPP 不同,其根本差别在于事先没有存入零件或零件族的典型工艺,零件的工艺过程主要是通过逻辑决策方式“创成”出来的。在研究创成式 CAPP 时,经常采用决策表和决策树两种方法。

1. 决策表

决策表是描述事件之间关系的一种表格,它可分为条件和决策两部分。决策表可用

图 6 .1 复合零件法示例

双线分成四个区域,其中左上区域是条件说明;右上角中各列是满足各种条件的组合,用 T 表示所在行的条件为真,空格表示“ 无关 ”;左下角是决策说明,即列出各种可能的决策行动,决策表的决策行动可以是无序的决策行动,用 X 表示,也可以是有序的决策行动。并给予一定序号,如表 6 .2 所示。

表 6 .2 决策表示例

尺寸精度 > 0 .1	T		
尺寸精度 < 0 .1		T	T
位置度 > 0 .1	T	T	
位置度 < 0 .1			T
钻 孔	X	1	1
铰 孔		2	
镗 孔			2

从本决策表上可看到,决策表中各条件的关系是“ 与 ”的关系,决策行动之间也是“ 与 ”的关系。决策表右面的每一列都可视为一条决策规则,如决策表右面第一行可写作“ 如果加工尺寸精度允差 > 0 .1,且位置度允差也大于 0 .1,则可采用钻孔加工。”

2 . 决策树

决策树是一种图,决策树有一个主根,并从此主根上分出多个分枝。如果某个分支的条件为真,则这一分支就表示一种可能的决策路线,每一个分支又可能通向某一中间结点

或终点。决策树作为一种常用的数据结构,也是一种决策逻辑的表达工具。它很容易与“如果(IF) .. 那么 (THEN) ...”这种直观的决策逻辑相对应,很容易转换成逻辑流程图和程序代码。图 6 .2 为表 6 .2 所示决策表的决策树。

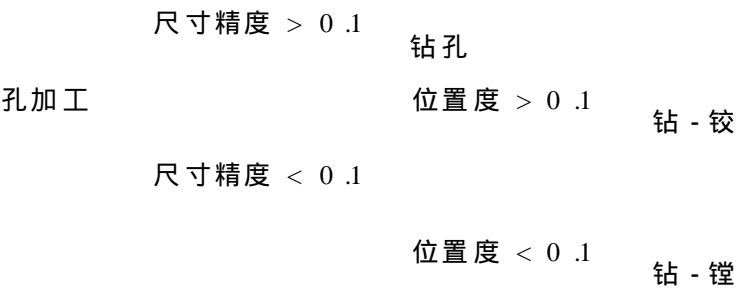


图 6 .2 决策树示例

6 .1 .6 CAPP 专家系统

随着人工智能技术的发展,专家系统已成为 CAPP 研究中的一个重要方面,从目前来看,专家系统在 CAPP 研究中的应用主要有以下两个方面;

(1) 专家系统用于从 CAD 模型中自动提取特征,从而为 CAPP 提供一种特征描述的方法。

(2) 开发 CAPP 专家系统。专家系统是一种工具,它具备某一领域的知识并利用这种知识来解决这一领域中的问题。采用专家系统的办法很适合于解决工艺设计中若干方面的问题。传统上,这些问题主要有两个主要特点:其一是工艺知识主要依靠经验,这些专门经验一般要通过较长时间的积累;其二是这些知识本身有一定的不确定性,且随企业条件的不同而变化。因此,如果利用专家系统来进行自动工艺设计,就必须和工艺师一样收集、表达和利用这些工艺知识,而专家系统本身为这些工作提供了很好的框架。

一个好的 CAPP 专家系统,应具备以下特点:

可维护性。近年来,我国的生产工艺处在不断发展阶段,新工艺、新设备和新材料发展很快,这就要求专家系统可不断补充和更新知识,以保证专家系统能不断适应改变了的情况。

开放性。一个好的 CAPP 专家系统应该有很好的开放性,以使用户能根据自己的情况对 CAPP 专家系统进行二次开发,也便于把该系统移植到类似企业。

良好的用户界面。系统应对用户有友好的用户界面,有很方便的输入和输出方式,向用户提出推理和决策的理由。必要时,工艺人员还可进行必要的人工干预,以制定出高质量的工艺规程。

6 .2 CAPP 系统中零件信息的描述

对于工艺设计来说,第一个任务就是读懂设计图纸。一个零件可以用多种不同的方式来表达,采用三向视图来表示所需加工零件的办法不但精确地表达了零件的几何形状和尺寸,而且在各种符号和文字说明的帮助下,工艺设计人员可以进一步准确无误地了解

到各种加工要求。

在进行一个零件的工艺设计时,工艺设计人员首先要了解零件的大体形状,如该零件为一回转体或箱体。然后进一步了解该零件的大小、加工特征、加工精度要求以及其它加工方面的信息。

在工艺设计过程中,我们把从设计信息转换为制造信息的过程称为设计工艺设计接口。这一部分工作可能是 CAPP 专家系统中最重要也是最困难的任务之一。

在 CAPP 专家系统的早期工作中,考虑到使用计算机把设计信息转换为制造信息的难度,设计接口(即读图)工作仍然由工艺人员来担任,即由人把工程图转换成 CAPP 系统可“理解”的数据格式,这种格式有时也被称为“特征”,然后采用交互方式输入计算机。对于比较复杂的零件,这种交互式的过程将变得十分繁琐而且容易出错,因此,有必要开展自动化的设计/工艺设计接口。

在自动化的设计/工艺设计接口的研究中,其中最重要的问题就是特征识别。在人工读图过程中,工艺人员首先做出一种假设,如工艺人员在工程图的某一个视图上看到一个圆,可能首先把它假设为一个孔,然后再到其它视图上收集信息来证实这一假设。如果这一假设失败,工艺人员可能把它假设为一个圆柱,在这一特征完全确认以后,可以把这一特征及其所包含的图素(点、线、圆等实体)从图上去除。这一过程不断进行,直到所有的特征都从上去除为止。当然,在实际读图时,人们并非真的把这些几何实体从图上擦掉。它们只是被收集拢来,以特征集合的形式存放在我们脑海中。

进一步说,自动化的设计/工艺设计接口问题的实质是从零件设计中提取加工特征的过程。近年来,在这方面已开展了大量研究工作并取得不少成果,但直到目前为止,还没有一种方法是尽善尽美的,而且离商业化应用还有一定距离。为了完成这一任务,还必须进行长期不懈的努力。

下面,我们将介绍几种常用的零件描述方法。

6 2 1 零件分类编码法

CAPP 中采用的零件分类编码可利用成组技术中比较成熟的编码方法,也可根据各个企业和产品的特点自行编制。当建立零件代码系统时,对编码的结构形式应根据实际需要加以选择,代码的结构一般可分为:

(1) 树式结构 树式结构也叫分级式结构,其特点是码位之间有隶属关系(级进关系),即后面的码位隶属于前面的码位。这种结构的每一个特征码位有很多分枝,很像树枝形状,所以称为树式结构。

(2) 链式结构 链式结构,也称并列结构。其特点是各码位之间不是从属关系,而是并列的。这种结构形式如同链条,故称链式结构。它所包含的信息量,在其它条件不变的情况下比树式结构少,但其结构简单、使用方便、应用较广。

(3) 混合式结构 混合式结构,即在系统中同时存在树式和链式两种结构形式。为了尽可能利用树式和链式两种结构形式的优点,扬长避短,目前大多数分类编码系统都采用混合式结构。

零件中包括的特征越多,即零件类型越复杂,描述这些零件特征的编码也就越长。我

们认为,在可能的条件下,代码应尽可能简短。这是因为过长的代码不仅使用不便,而且不论代码多长,都不可能详尽地描述出零件的全部信息。

目前,国内外已有几十种,甚至上百种编码方法,其中奥匹兹(OPITZ)分类编码系统是在国际上影响较大,也是较成熟的一种系统。

奥匹兹系统采用 9 位数字码排列组合描述零件,前 5 位用来描述零件的形状和结构特征,称为主码(又称形状码);后 4 位分别表示零件的尺寸、材料、原始形状和精度,称为辅码。每个码位内有 10 项,分别表示零件的 10 种特性。

我国自行研制的机械零件分类编码系统 JLBM-1(JB/ Z251-85)是一种比较适合我国机械行业的成组编码系统。该系统共有 15 个码位,其中第 1、第 2 码位为零件名称类别矩阵(用两个码位反映零件的类别和名称),第 2 码位主要描述零件功能,如螺母、链轮、压板和凸轮等。第 3 码位~第 9 码位是零件的形状和加工码位,在 JLBM-1 系统中采用把回转体和非回转体零件分开描述的办法。分开描述的好处是使回转体零件和非回转体零件的加工信息都可得到比较充分的描述。10 码位,11 码位,12 码位分别表示材料、毛坯原始形式和热处理。13 码位和 14 码位用于描述零件的尺寸,并按尺寸档次分开描述大、中、小件。这样做的好处是不仅对零件尺寸描述得更加细致,而且扩大了系统的应用范围。第 15 码位主要用来描述零件的精度,它把加工精度分为低、中、高和超高精度 4 级,比较符合我国实际情况。

成组编码的主要原理是相似性原理,按成组原理对零件进行分类的办法已广泛应用于派生式 CAPP 系统。按成组技术思想对于零件的加工工艺规则进行分析和总结,其中许多规则也可用于创成式 CAPP 系统或进入 CAPP 专家系统的知识库。

6 2 2 零件的图形要素描述法

回转体零件的描述方法比较多,有许多已成功地应用于各种 CAPP 专家系统,下面介绍的只是许多描述方法中比较有代表性的一种。

任何一个零件都是由一些有关的表面相互连接而成的,这些表面可以分为主表面和辅助表面等,对于回转体零件,则可以进行如下划分:

1. 信息代码的规定

(1) 主表面	代码
端面	DM 或 dm
圆柱面	YZ 或 yz
圆锥面	ZM 或 zm
螺纹面	LW 或 lw
花键面	HJ 或 hj
齿轮面	CL 或 cl
棱柱面	LZ 或 lz
球面	QM 或 qm
圆环面	YH 或 yh

(2) 辅助表面	代码	副码			
倒角	DJ	L: 左倒角	R: 右倒角		
倒圆	DY	L: 左倒圆	R: 右倒圆		
键槽	JC	1: 平键	2: 半圆键		
圆周孔	YK	1: 单孔	2: 均布孔	3: 非均布孔	
圆周槽	YC	1: 环形槽	2: 圆弧槽	3: V 形槽	4: 越程槽
中心孔	ZK	1: 60 度	2: 120 度	3: 螺纹中心孔	
端面孔	DK	1: 单孔	2: 均布孔	3: 非均布孔	
过渡圆角	YJ				
滚花	GH				
孔型		1: 光孔	2: 螺纹孔	3: 锥孔	4: 台阶孔

(3) 形状公差	代码	形状公差	代码
直线度	ZX	圆柱度	ZD
线轮廓度	XK	面轮廓度	MK
平面度	PM	圆度	YD
(4) 位置公差	代码	位置公差	代码
平行度	PX	垂直度	CZ
倾斜度	QX	同轴度	TZ
对称度	DC	位置度	WZ
圆跳动	YT	全跳动	QT

(5) 定形尺寸

- A 角度(如锥度、斜孔的角度等)
- B 宽度(如槽宽、齿宽)
- D 直径或大径(大径指矩形花键小径和圆锥面大端直径)
- Da 齿顶圆直径
- d 小径(小径指矩形花键小径和圆锥面小端直径)
- H 槽高(指槽底或单一平面到圆柱体底边回转线距离)
- L 长度(如圆柱面、锥面、螺纹、花键和槽的长度)
- M 螺纹直径
- N 正棱柱的边数或花键的齿数
- R 半径(如圆弧、球面、圆环、倒圆、键槽铣刀半径等)
- S 正棱柱边长

(6) 定位尺寸

- LL 某表面元素到左端面的距离
- RL 某表面元素到右端面的距离
- JL 某表面元素到基准面的距离
- DD 端面孔或槽的中心线直径
- RR 端面孔或槽的中心线半径
- JD 某表面元素到基准的旋转角度(以圆柱左视 X 正向为 0 度)

2. 表面元素的输入原则

零件各表面信息按照主轴线顺序进行输入,先外后内,外表面从左到右,内表面从右到左依次进行,其语法规则如下:

序号 主表面代码[(定形尺寸 [,表面粗糙度][,形状公差][,位置公差][,定位尺寸])]

[{辅助表面代码[(定形尺寸 [,表面粗糙度][,形状公差][,位置公差][,定位尺寸])]}]

其中,方括号中为可选参数,尖括号中为必选参数,说明如下:

对于主表面,需先输序号,然后输主表面代码,接着在圆括号中输入主表面本身的有关信息。如该主表面包含辅助表面,则在花括号中输入辅助表面代码,接着在圆括号中输入辅助表面的有关信息;若辅助表面还包含辅助表面,则依次类推。

有精度的尺寸在尺寸数字后的圆括弧中输入其上下偏差值。

3. 表面信息输入方法

(1) 端面(DM 或 dm)

对于回转体零件,其端面基本上为回转面所包容的圆平面(只有一个外径)和圆环面(一个外径和一个内径)。由于这些尺寸与相应的回转面尺寸相关联,在回转面输入时要加以输入,为避免重复,这里就不再输入。

如图 6.3 中表面 1,直径为 55(0.023,0.003),表面粗糙度为 $Ra = 6.3$,可表示为:

DM($Ra\ 6.3$)

(2) 圆柱面(YZ 或 yz)

圆柱面的定形尺寸是直径 D 和长度 L,这两个尺寸是必选参数,一定要输入,如图 6.3 中的表面 2 可表示为:

YZ(D 55(0.023,0.003),L 70, $Ra\ 1.6$)

(3) 圆锥面(ZM 或 zm)

圆锥面有三个尺寸,可标注小端直径 d,大端直径 D 和锥体长度 L,也可标注小端直径 d,大端直径 D 和角度 A,输入时先输左端直径,后输右端直径,例如:

ZM(D60,d30,L20, $Ra\ 3.2$)表示大端直径为 60,小端直径为 30,锥体长 20,表面粗糙度 $Ra = 3.2$ 的左端大,右端小的外锥面。

(4) 倒角(DJ)

倒角有两个尺寸,一个是长度,一个是角度,按图纸习惯标注,不需另加符号,倒角处于圆柱面的两头,应用副码指出方向,例如:

DJL(2 * 45),表示圆柱面左端 2 * 45 度的倒角。

(5) 键槽(JC)

键槽可分为平键和半圆键,需用副码加以区别,下面以平键为例加以说明:

对于平键,除了要标出槽长 L,槽宽 B 和槽高 H 外,还要确定键槽的定位尺寸,即槽到某端面的距离,LL 表示离左端面的距离,RL 表示离右端面的距离,还应指出槽在回转

面上的角度 JD, 例如图 6.3 表面 2 上的键槽:

JC1(L50,B16(0, - 0.05,Ra 6.3),H50(0, - 0.17),DC(0.08|4),LL10,JD0)

(6) 端面孔

端面孔指在端面上分布的各种孔, 举例如下:

DK21-4(D10,L0,Ra 6.3,DD140,JD0)表示 4 个均布光孔, 孔径为 10, 孔的中心线直径为 140, L0 表示通孔, 孔的表面粗糙度为 Ra=6.3。

4. 实例

下面以图 6.3 为例, 说明一个具体零件信息代码的输入方法。

图 6.3 实例

序号	信息内容
1	DM(Ra 6.3)
2	YZ(D 55(0.023,0.003),L 70,Ra 1.6) {DJL(2 * 45),JC1(L50,B16(0, - 0.05,Ra 6.3),H50(0, - 0.17), DC(0.08 4),LL10,JD0)}
3	DM(Ra 6.3)
4	YZ(D 75,L70,Ra 1.6)
5	DM(Ra 6.3)
6	YZ(D176,L30,Ra 1.6)
7	DM(Ra 6.3) {DK21 - 4(D10,L0,Ra 6.3,DD140,JD0)}

6 2 3 直接从 CAD 中提取信息

CAD 系统中已包含了有关零件在形状、结构、尺寸等多方面的详细信息,对这些信息进行分析、识别、转换,进而提取出 CAPP 系统所需要的工艺信息,这是一种理想的方式。国内外近年来对此开展了大量的研究工作,特别是在加工特征提取方面取得了不少研究成果。

加工特征是一组与加工有关的几何实体,例如通孔、盲孔、凸台、台阶、键槽、平面、曲面、螺纹孔等。在加工特征和加工工艺之间有松散的对对应关系,如箱体零件上的螺孔一般可通过钻孔、攻丝来加工。如果识别出来的特征为一键槽,则可采用铣削加工。

对于工艺人员来说,确定零件的几何特征可能是一件容易的事,但对于计算机程序,就是一件不十分容易的事了。采用专家系统来识别几何特征就是办法之一。

首先来看如图 6.4 所示的台阶和槽的特征。

图 6.4 台阶和槽

下面是识别上述几何特征的产生式规则:

RULE 1 (台阶识别)

IF face f_1 is adjacent to face f_2

AND the angle between face f_1 and f_2 is 90°

THEN the faces f_1 and f_2 form a step

规则 1 (台阶识别)

如果面 f_1 和 f_2 相邻

且面 f_1 和 f_2 之间的夹角为 90°

那么面 f_1 和 f_2 构成一个台阶。

RULE 2 (直角槽识别)

IF face f_1 is adjacent to f_2

AND face f_2 is adjacent to f_3

AND the angle between f_1 and f_2 is 90°

AND the angle between f_2 and f_3 is 90°

THEN the faces f_1 , f_2 and f_3 form a slot

规则 2 (直角槽识别)

如果面 f_1 和 f_2 相邻

且面 f_2 和 f_3 相邻

且面 f_1 和 f_2 之间的夹角为 90°

且面 f_2 和 f_3 之间的夹角为 90°

那么面 f_1 , f_2 和 f_3 构成一个槽。

用产生式规则确定全部零件特征的搜索过程,其计算比较复杂。利用一种相邻表面属性图(AAG)和正向推理方法来识别特征,可大大降低计算开销。AAG 可以被定义为一种图 $G=(N, A, T)$, 其中 N 为结点的集合, A 为弧的集合, T 为弧 A 的属性的集合。

在 AAG 中,对应于零件上的每一个面,在集合 N 中都有一个唯一的结点 n 与其对应。相邻两个面之间的关系用一条弧 a 来表示。如果相邻两个面所构成的角是凸的,那么属性值 a 为 1;如果相邻两个面所构成的角是凹的,那么,其属性值为 0。这样,特征识别问题就转化为一个 NP 问题,由于目前尚无一种理想的算法,当结点数变得很大时,计算时间将会大得无法承受。

采用 AAG 方法,图 6.4 中的台阶和槽可用以下规则来识别:

RULE 3 (台阶识别)

IF graph is linear, and

has exactly two nodes with an incident arc with attribute 0

THEN

feature is a step

规则 3 (台阶识别)

如果 图为一 直线,且

只有两个结点,连接这两个结点的弧的特性为 0

那么

该特征为一台阶。

图 6.5 为一台阶的 AAG 图。

图 6.5 台阶的 AAG 图

采用上述办法,可以为每个特征编写一条规则。

在开始搜索特征子图以前,首先要进行结点删除步骤,这一步骤删除这样一些结点:连接这些结点所有的弧的属性都是“1”。这样就可把一个相邻表面属性图拆成几个互不相连的子图。把这一方法用于某一零件的 AAG(图 6.6(a))可得出如图 6.6(b)所示的结果,图中每一个子图表示一个特征或几个相交的特征。

采用一种匹配过程可识别单一的特征。几个相交的特征可能共有一些表面,因此,识别它们就需要更多一些步骤。两个相交的特征可用结点分离法将它们分离开来。例如,在图 6.7 中有两个相交的槽,其中表面 1 为两个槽所共有,而表面 2 和表面 3 原来是同一表面。结点分离程序把表面(结点)1 分成两个表面,而结点合并程序可把表面 2 和表面 3

图 6.6 某一零件的 AAG

结合成表面 2/3。这一方法比较适合二维半的特征。通过以上简化办法可省略许多不必要的搜索,然而,上述办法还需要进一步的细分类。例如,我们可以从一个零件中识别一个槽,然后根据槽的底角,进一步把槽细分为直角槽和燕尾槽两种。

图 6.7 带有两相交槽的零件

国外在特征提取方面已做了大量的研究工作,据文献资料介绍,目前已有不少 CAPP 原型系统采用了特征提取的办法从 CAD 中提取信息。

6.2.4 特征建模

CAD 系统应成为 CAPP 系统的数据源,CAPP 需要的零件描述信息包括了零件的整体形状、加工部分的形位公差、表面粗糙度等高层信息。这恰是传统 CAD 系统的弱点所在。目前市场上销售的三维造型 CAD 软件,如 Euclid, Unigraphics, I-deas, CADDs, Pro/ Engineer 等,虽然都带有“特征造型”模块,但其信息中只包括零件的几何形状和拓扑结构,造型数据同样无法描述完整的产品信息。如公差、表面粗糙度等,这使它们也无法彻底摆脱传统 CAD 系统的局限性,不能强有力地支持工程应用所需要的高层信息。而且对用户而言,大多数商品化软件为了技术保密需要,数据结构是封闭的,更给用户实现 CAD/CAPP 的集成造成了极大的困难。

特征造型也可称为基于特征的设计,它是一种产品建模的方法,其主要特点是把几何和非几何信息全部汇入产品(零件)定义中,因而是一种理想的造型方法,也是产品造型的

发展方向。在本书第 4 章中已对特征建模作了比较详细的介绍,因此这里不再重复。

STEP 标准是国际标准化组织提出的一个产品模型数据交换标准,它是一个中性的产品模型数据交换机制,表示了贯穿产品生命周期的产品定义数据,并为各个计算机辅助工程应用系统之间的数据交换提供通道。STEP 包含三方面重要内容:一是参考模型(reference models),为进行完整的和无二义性的产品描述提供必要的产品定义模式;二是 EXPRESS 形式化语言,它是用来进行描述产品的数据、定义数据结构、操作和约束的计算机语言;三是 STEP 文件结构(file structure),为数据通信和取用提供一个有效和可靠的模式。

CAPP 是产品开发生命周期中的一个重要环节,在目前发布的 STEP 参考模型中,与 CAPP 应用系统有关的有 4 个参考模型:公称形状信息模型(nominal shape information model),它表示零件的公称形状,包括几何、拓扑与实体(Solids)等;形状特征信息模型(form feature information model),定义了具有特定形状的特征;形状公差模型(shape tolerance model),定义了由相关 ISO 标准给出的尺寸公差信息;表面信息模型(surface information model),定义了表面粗糙度、表面硬度方面的信息。

应该说,开发基于 STEP 的新一代的 CAD/ CAPP 系统是最理想的途径。可是由于 STEP 本身尚在逐渐充实与成熟的阶段,基本框架与方法虽已提出,具体内容却需继续研究与补充,而且全新系统的商品化过程耗时较长,投资也大。因此,利用现有的商品化 CAD 系统,扩充特征造型功能,使产生的 CAD 模型能表达必要的工艺信息,并与 CAPP 需要的信息格式实现一致的特征造型已成为我国 CAPP 研究的热点之一。

6.3 工艺知识的获取与处理

6.3.1 工艺知识的获取

在 CAPP 专家系统中,由于工艺知识是建造专家系统的基础,因而在一个 CAPP 专家系统的建造中,解决知识如何获取是一个十分重要的课题。在专家系统的开发过程中,人们发现,知识获取是最难解决的一环,被认为是专家系统建造中的“瓶颈”问题。大量的事实表明,知识获取的过程消耗人力、物力最多,所用的时间也最长。据统计,一般要占整个专家系统开发中总工作量的四分之三。工艺经验知识是一种领域(专业)知识,工艺知识的获取途径很多,概括起来主要有以下几个方面:

(1) 机械制造工艺学及有关文献 机械制造工艺学提供了机械加工工艺的基本原理和基本概念,这些原理和概念是指导工艺设计的基本准则。在收集有关工艺知识时,对企业现行工艺违反工艺学基本理论的地方应予以纠正,避免一些错误的知识存入知识库。

(2) 企业现行工艺文件 企业现行工艺文件是建造专家系统所需要的最主要知识源,这是因为企业工艺文件一般都经过长期的生产实践考验,被证明是符合本企业实际情况而且是行之有效的。工艺文件反映了工艺设计中所需要的工艺知识,可以通过工艺文件分析工艺设计所需要的规则和各种工艺知识间的逻辑关系。

(3) 各种有关的规范、标准和手册 根据系统功能的要求,确定并选择有关的规

范、标准和手册,如机械加工工艺师手册等。

(4) 从有经验的工艺人员和工人技师中获取知识 选择具有较高水平的工艺人员和工人技师作为调查对象,可以请他们作专题介绍,也可请他们解答一些具体工艺问题。

知识源提供的知识是未经处理的原始型知识,将其输入知识库之前,必须采用适当方法对其进行处理、解释、挑选、加工和改造,以便形成规律、建立规则、消除冗余,排解矛盾,使工艺知识统一化、规范化、规格化。

知识的获取和处理是一项十分困难的工作,其难点主要表现在以下几个方面:

(1) 从领域专家获取知识是一件艰巨的工作 工艺知识的获取必须在知识工程师(软件开发人员)和领域专家(工艺人员和工人技师)两者密切下合作共同努力实现。一般来说,领域专家可以很快地编制出一个零件的加工工艺,但他们却往往不善于或者说不出为什么要这样做。他们往往不能将这些知识总结成规律性的结论。另一个知识获取的困难是多个领域专家的知识之间相互矛盾,如何处理这些矛盾,也是知识获取中的难点之一。

(2) 对工艺文件知识处理的困难 工艺文件是获取知识的最主要的知识源。然而工艺文件上反映的知识是工艺人员对某个零件加工进行一系列决策的结果,而不是决策的过程和依据,因而是不完全的工艺知识。又由于工艺人员的水平有高低之分,工艺的编制时间和当时的设备条件也有所不同,因而编制的工艺存在着多样性。这就造成了工艺文件上反映的知识往往存在离散性、随机性和模糊性。因此,通过工艺文件去揭示工艺设计的原理、方法和各种工艺规则,以及各工艺知识间的逻辑关系,自然是一件十分困难的事。

鉴于上述困难,我们必须重视知识的获取和知识的处理方法的研究,下面我们将简要介绍一下工艺知识获取的基本方法。

1. 分类归纳法

对从知识源获得的原始知识进行处理的目的是从表面上看是离散的、杂乱的、模糊的工艺知识,经过去粗取精,去伪存真的过程,从中找出工艺设计的规律,建立起一条条工艺规则。传统的工艺设计是针对一个个零件单独进行的,它容易忽视同类零件的联系和内在的规律。利用成组工艺的原理,将零件分类成组,把同组零件的工艺集中起来,对其进行归纳分析,这样就比较容易找出规律性的工艺知识,也更容易发现组内工艺知识存在的变异。

例如,有三个圆柱直齿轮,它们的工艺路线分别为:

齿轮 $A = \{\text{粗车, 正火, 半精车, 平磨, 滚齿, 倒齿角, 去毛刺, 剃齿, 热处理, 磨孔, 拉键槽, 去毛刺, 珩齿}\}$ 。

齿轮 $B = \{\text{粗车, 正火, 半精车, 平磨, 滚齿, 去毛刺, 剃齿, 热处理, 推孔, 拉键槽, 去毛刺, 珩齿}\}$ 。

齿轮 $C = \{\text{粗车, 正火, 半精车, 平磨, 滚齿, 去毛刺, 剃齿, 热处理, 磨孔, 拉键槽, 去毛刺, 珩齿}\}$ 。

经分析,可找出其共有的工序为:

齿轮 A = {粗车, 正火, 半精车, 平磨, 滚齿, 去毛刺, 剃齿, 热处理, 磨孔, 拉键槽, 去毛刺, 珩齿}。

其共有工序表明, 加工反映了这一组齿轮精度要求应遵循的共同规律, 是形成该类零件工序规则的依据。

组内部分零件共有或单个零件独有的工序形成了与共有工序的变异, 产生工艺变异的原因有两种: 一种是合理变异, 因为零件的设计或加工要求不同而必然产生的变异, 如齿轮 A 中的有倒齿角工序, 就是齿型要求造成的。对合理变异处理方法是增加附加规则。另一种是不合理变异, 是由于编制工艺缺少统一规范, 由工艺师个人决策造成的。上述三个齿轮的加工, 齿轮 A 为磨孔 (直径为 42mm, 精度 H7), 齿轮 B 为推孔 (直径为 25mm, 精度 H7), 齿轮 C 为磨孔 (直径为 25mm, 精度 H7)。齿轮 B 和齿轮 C 孔径相同, 加工精度要求也相同, 但加工方法不同, 这是不合理变异, 对不合理变异应在知识处理阶段消除。根据工艺规程, 在零件直径大于 25mm 时应采用磨孔, 零件直径小于或等于 25mm 时采用推孔, 因此, 齿轮 A 的磨孔和齿轮 B 的推孔符合规律, 齿轮 C 应改为推孔。

经过上述分析后, 该组齿轮的加工工序在共同工序规则的基础上, 应附加以下三条规则:

- 如果齿形有倒角要求, 则增加倒齿角工序;
- 如果齿轮内孔直径大于 25mm, 则采用磨孔工序;
- 如果齿轮内孔直径小于或等于 25mm, 则采用推孔工序。

分类归纳法适用于对工艺路线 (包括工序及其顺序, 工步及其顺序) 的分析。

2. 统计分析法

由于传统工艺设计造成的工艺多样性的存在, 因此工艺卡片上反映的工艺参数或工艺常识常呈分散状态。用统计分析法可发现工艺知识在数量上的分布规律, 根据分布规律形成相应的工艺规则, 根据分布规则可形成相应的工艺规则。统计分析法适用于加工方法、加工手段 (机床、刀、夹、量、辅具) 和某些工艺参数的加工处理。

例如用统计分析法确定齿轮类零件 7 级孔的加工方法, 首先可将有关零件的工艺卡片集中, 将孔的有关参数与加工方法按一定格式记录下来, 如表 6 3 所示, 然后对数据进行统计分析得出结论。

表 6 3 齿轮类孔加工统计表

孔径 (mm)	粗车工序	半精车工序			推孔	磨孔	拉孔	零件号
	(钻)	(钻)	(扩、镗)	(铰)				
10H7		9 8		10H7				
12H7	10		11 8	12H7				
18H7	16		17 8	18H7				
25H7	23		24 8	25H7				
25H7	23		24 8	25H7				

续表

孔径 (mm)	粗车工序	半精车工序			推孔	磨孔	拉孔	零件号
	(钻)	(钻)	(扩、镗)	(铰)				
25H7	23		24	24 3			25H7	
35H7	33		34 4	34 .7		35H7		
40H7	38		39 .7	40H7				
40H7	38		39 4	39 .7		40H7		
42H7	40		41 .7	42H7		(校磨)		
45H7	43		44 2	44 5		45H7		
47H7	45		46 2	46 5		47H7		
...		

由上述数据可以看出如下规律：

(1) $D < 10\text{mm}$ 时,加工方法为：

粗车：不安排钻孔

半精车：钻孔 ($D - 0.2$), 铰孔 DH7

(2) $10 < D < 25\text{mm}$ 时：

粗车：钻孔 ($D - 2$)

半精车：扩(镗) ($D - 0.2$), 铰孔 DH7

推孔

(3) $25 < D < 60\text{mm}$ 时

粗车：钻孔 ($D - 2$)

半精车：扩(镗) ($D - 0.6$), 铰孔 ($D - 0.3$)

$D < 40\text{mm}$

扩(镗) ($D - 0.8$), 铰孔 ($D - 0.5$)

$D > 40\text{mm}$

磨孔：磨孔 DH7

3. 数学模型决策法

在工艺设计中有些参数可以通过建立数学模型进行决策时,应尽量建立数学模型。有许多工艺知识,如尺寸链的计算、工时定额、材料定额等都可以采用这一办法。

6.3.2 工艺知识的表达

工艺设计的主要目的是把设计转化为加工工艺,为实现这一目的,CAPP 系统必须具备有关加工工艺方面的知识。这些加工工艺方面的知识可以用多种方式来表达,如决策树、决策表和专家系统规则等。在本节中,我们将主要讨论一些适合于 CAPP 专家系统的工艺知识的表达方法。

工艺知识库是 CAPP 专家系统的主要组成部分,生产实际中实用的 CAPP 专家系统所需要的信息量是很大的。在 CAPP 专家系统的工艺知识库中,一般至少应包含以下几方面的规则:

- 毛坯选择规则;
- 加工方法选择规则;
- 排序规则;
- 定位装夹规则;
- 设备选择规则;
- 刀具、量具、辅具选择规则。

为了使知识基系统的知识库更有条理,同时也便于以后的维护和改进,大多数的知识库都按知识性质细分为许多子库,如加工方法选择子库,设备选择规则子库,刀具、量具、辅具选择规则子库等。在某些系统中还把这些子库再细分成若干子库,使每个子库对应着工艺决策的一个相应子任务。

工艺知识的表达方式很多,在专家系统中常用的一些知识表达方式有谓词逻辑、产生式规则、语义网络、框架、剧本等。根据发表的文献资料来看,在 CAPP 专家系统中使用得最多的知识表达方式是产生式规则,其次是框架模型。近年来又加强了对一些新的表达方法或原有方法的综合应用研究,如面向对象的方法、混合式知识表示模式以及各种模糊知识表示等。下面我们举例说明其中几种表达方式。

1. 产生式规则

(1) 加工方法选择规则 加工方法选择的目的是确定各特征表面的加工链。举例如下:

```
IF
    (加工对象:“ 圆柱面 | 圆锥面 ”)
    (工件材料: 普通碳钢 | 优质碳钢 | 合金钢)
    (表面淬火最低硬度 HRC:“ 38: none ”)
    (最高表面粗糙度 Rz:“ 1.6 ”)
    (最高加工精度 IT:“ 7 ”)
THEN
    (推荐采用:“ 粗车, 精车, 磨削 ”)
```

(2) 排序规则 排序的目的是将已选好的各种特征表面的加工方法合理安排,以生成工艺路线。举例如下:

```
IF
    (加工零件为:“ 实心轴 & 单向阶梯 & 两端有中心孔 ”)
THEN
    (加工顺序为:“ 小头端面, 中心孔, 小头各外圆, 调头, 大头端面, 中心孔, 大头各外圆 ”)
```

2 . 工艺知识的框架表示

框架结构适宜于表示知识的层次结构,提供有效组织知识的手段,易于处理默认值并能协调叙述性知识与过程性知识的相互关系,在工艺 CAPP 专家系统中得到比较广泛的应用。

下面是描述零件加工工艺流程的一种框架形式:

```
( 工艺流程名称  工序 1 ( 装夹方式  文字 )
    ( 夹具名称  文字 )
    ( 工步 1    加工方法  文字 )
        ( 加工表面  文字 )
        ( 刀具名称  文字 )
        ( 切削参数 1  字符 )
        ...
    ( 切削时间  字符 )
    ...
( 工序 2 ... )
...
... )
```

3 . 基于对象的工艺知识表示模式

下面举例说明一种基于对象的工艺知识表示模式,在这种表示模式中,对象类的总体描述为:

对象类名 , 超类 , 属性 , 约束 , 方法 共五部分组成。其中:

“对象类名”描述了对本对象类的名称标识,“超类”记录了该对象的父辈对象,“属性”描述了对对象的属性的名称、值型等,“约束”描述了对属性间的联系及属性值满足的条件,“方法”记录了对该对象的操作及每种操作所遵循的规则集。

在对象描述中,每一个约束可作为一个规则元。规则元是具有确定信息的一个判断、关系表达式或指令(过程或赋值),它是组成知识的基本元素。

一个规则元可总体描述为:

左项 , 操作符 , 右项

共三个部分组成。其中:

“左项”及“右项”可以为常量、字符串、变量、表达式或函数、命令等,“操作符”描述左项、右项之间的关系或操作。

引入对象和规则元概念以后,可以这样定义规则:

规则用来描述对象之间的相互联系,是由多个规则元以一定方式和顺序(IF-THEN-ELSE 形式)组织关联在一起的信息实体。

一条规则可以总体描述为:

条件(IF)规则元序列 , 结论(THEN)规则元序列 , 否则(ELSE)规则元序列 , 权重 。

6.4 应用实例——回转体零件 CAPP 专家系统中的夹具设计

在计算机辅助工艺设计中,夹具设计是很重要的一环,但是在已有的许多 CAPP 专家系统中,这一点往往被忽略。因为夹具选择的正确与否将直接影响到产品的质量和零件加工所需的辅助时间的长短,而缩短辅助时间对于提高机床的利用率以及提高产量,降低成本有很重要的意义。

因为对于回转体零件的车削加工来说,大多数夹具都是现有的,所以回转体零件 CAPP 专家系统中的夹具设计实际上是一个选用最优夹具的问题。在 CAPP 专家系统的夹具设计中,以下几方面的信息是必不可少的:

- 零件的几何、拓扑信息;
- 零件的形位、公差信息;
- 某给定表面的加工方式;
- 加工工艺参数(切削速度、切削深度、进给量);
- 现有夹具明细表;
- 每一种夹具的特征。

本回转体零件 CAPP 专家系统中的夹具设计模块,必要时也可作为一个单独系统附加在已有的回转体零件 CAPP 专家系统后面,起一种“后处理”的作用。程序采用人工智能语言 Prolog 编写。

6.4.1 夹具设计模块的数据库和规则库

可以把回转体零件夹具选择的规则分为以下三类:

(1) 定性规则 例如,如果被夹持表面的截面为六边形,则不应采用四爪卡盘。在工件无通孔的情况下才有可能用两顶尖夹紧。

(2) 定量规则 定量规则与加工和零件参数有关,例如,如果某种装夹方式造成零件的加工误差(由于切削力、工件悬臂过长等原因)超过允差范围,那么,不能采纳这种装夹方式。

(3) 启发式规则 例如,如果采用两顶尖夹持方式,且工件的长度/直径比在 6~10 之间,则应采用一个固定中心架。

在进行夹具选择的决策过程中,还需要有以下数据库和文件:

- 制造数据库;
- 零件的特征信息文件;
- 工艺信息文件;
- 夹具信息文件。

其中制造数据库是一个通用数据库,它与工件和车间的具体情况无关,零件的特征信息文件和工艺过程文件的内容主要取决于工件。夹具信息文件则和车间的具体情况有关。下面我们将详细介绍以上所述的数据库和文件。

1 . 制造数据库

在制造数据库中,各种有关制造信息,如各种不同材料的信息都以谓词逻辑表示,在 Prolog 中它们以子句的形式出现。例如,以“材料”命名的子句形式如下:

```
material(name of the material,  
         specification of the material,  
         specific horsepower of the material,  
         modulus of elasticity of the material) .
```

材料(材料名,
 代号,
 单位切削量消耗功率,
 材料的弹性模量) .

在本制造数据库中,只存放与夹具设计有关的制造信息。

2 . 特征信息文件

在利用 CAD 进行零件设计以后,可利用特征提取方法生成该零件的特征文件 (Feature Information File),在零件的特征文件中,零件的各个表面的信息也是用谓词逻辑来表达。其子句格式如下:

```
feature(feature number,  
        surface type,  
        surface location,  
        starting point x coordinate,  
        starting point y coordinate,  
        end point x coordinate,  
        end point y coordinate,  
        length of the surface,  
        diameter of the surface,  
        height of the surface,  
        information on whether machining is required,  
        the tolerance associated with the surface,  
        the cross section of the surface) .
```

特征(特征号,
 表面类型,
 表面位置,
 起点坐标 X,
 起点坐标 Y,
 终点坐标 X,
 终点坐标 Y,

表面长度,
表面直径,
表面高度,
是否需要加工,
与表面有关的公差,
截面形状) .

因为每一个子句描述一个特征(或表面),因此,一个有 11 个特征的表面就需要一个具有 11 行子句的特征文件来进行描述。一个零件可以包含许多特征,特征文件名由零件名 + 括展名(FIF)组成。

3. 工艺信息文件

本文件主要用来存放已生成的零件加工工艺信息,它可由 CAPP 软件包的其它模块生成。和上述特征信息文件一样,加工工艺信息也用子句来表达,其格式为:

```
operation( peration number ,  
           operation type ,  
           operation description ,  
           list of surface to be machined during the operation ,  
           speed ,  
           feed ,  
           depth of cut ) .
```

操作(操作号,
操作类型,
操作描述,
在本操作中需加工的表面,
切削速度,
进给量,
切削深度) .

下面是一个实际例子:

```
operation(8, turn , rough , [5,6] ,100,0 .1,1 .2) .
```

该子句描述了工艺设计中一个操作号为 8 的操作,操作类型为粗车,在这次操作中待加工的表面为表面 5 和 6,切削速度为 100m/ min,进给量为 0 .1mm/ r,切削深度为 1 .2mm。

4. 夹具信息文件

这一文件给出了某一特定环境下(例如某一车间)所能提供的所有夹具。每一种类型的夹具都采用独立的谓词逻辑表示,典型的谓词及其相应的内容表示如下:

```
three paw chunk (brand of the chuck,
```

maximum outside diameter that can be held by the chuck ,
smallest outside diameter that can be held by the chuck ,
maximum inside diameter that can be held by the chuck ,
smallest inside diameter that can be held by the chuck ,
the length of each chuck paw ,
the rated speed of the chuck) .

三爪卡盘(卡盘牌号 ,
卡盘上能夹持工件的最大直径(外夹) ,
卡盘上能夹持工件的最小直径(外夹) ,
卡盘上能夹持工件的最大直径(内夹) ,
卡盘上能夹持工件的最小直径(内夹) ,
夹爪长度 ,
卡盘额定转速) .

collet (rand of the collet ,
maximum outside diameter that can be held by the collet ,
minimum outside diameter that can be held by the collet) .

弹簧夹头(夹头牌号 ,
夹头能夹持工件的最大外径 ,
夹头能夹持工件的最小外径) .

mandrel(ummy variable since the brand is unimportant to decision ,
minimum diameter of hole that can be used as holding surface ,
maximum diameter of hole that can be used as holding surface) .

...

心轴(因为牌号对于决策不很重要所以采用哑变量 ,
夹持表面的最小孔径 ,
夹持表面的最大孔径) .

...

machine tool(brand of machine tool ,
maximum diameter of workpiece over way-swing of lath ,
machine tool error specified by manufacturer) .

机床(机床牌号 ,
床身上最大回转直径 ,
厂方提供的机床误差) .

在本程序中,规则同样用子句来表达。首先,让我们来看一下谓词“夹具”。谓词“夹具”有以下对象:

fixture(operation number ,fixture name, brand, holding surface number)

夹具(操作号, 夹具名称, 牌号, 夹持表面号)

因为操作号和夹持表面号可由 CAPP 其它模块提供, 所以这两个参数已被约束。这样, 如果对某一特定的夹具, 比如说名为“三爪卡盘”的夹具, 其牌号为“长城”, 那么, 规则可用下述形式表示:

```
fixture(operation # , three pawchuck, greatwall, holding surface # ):-
```

```
Rule 1 ...          and
```

```
Rule 2 ...          and
```

```
...                and
```

```
Rule (n - 1) ...    and
```

```
Rule n ...
```

式中:-在 Prolog 语言中表示“如果”, 也可用 if 代替。Prolog 的规则由三部分组成, 头、体以及隔开两者的符号:-, 跟在头和:-符号后面的是规则体, 规则体包含一个或多个子目标, 为了使规则成功, Prolog 必须满足其所有子目标。

夹具(操作号, 三爪卡盘, 长城, 夹持表面号):-

```
子目标 1 ...      且
```

```
子目标 2 ...      且
```

```
...              且
```

```
子目标 (n - 1) ... 且
```

```
子目标 n ...
```

对于某一定的操作号和夹持表面, 如果子目标 1, 子目标 2..., 子目标 n 都被满足, 那么该规则成功。不但如此, Prolog 程序还将通过回溯找出所有可能的解。事实上, 对于某一定的操作号和夹持表面, 往往不止一种夹具可选用。

现在, 可以给定一个目标, 其操作号和夹持表面已被约束为某指定值, 而夹具名称和牌号为变量, 例如, 如果系统给出一个目标: 找出所有能用于操作号为 10, 夹持表面号为数 2 的夹具。这一目标可表达为以下形式(注意在 Prolog 中变量的第一个字母必须大写):

```
fixture(10, Fixture- name, Brand, 2),
```

夹具(10, 夹具名, 牌号, 2) .

借助 Prolog 的合一功能, 系统将寻找实现这一目标的所有可能的解, Fixture- name 和 Brand 将被赋值。

有时夹具需要一些附件, 比如跟刀架, 在车削细长工件时被用来减少工件变形, 从而降低工件的加工误差。谓词“fix- accessories”可定义如下:

```
fix- accessories(operation number,
```

```
    fixture name,
```

```
    hold surface,
```

```
    name of fixture accessory) .
```

夹具附件(操作号,

夹具名称，
夹持表面，
夹具附件名称)。

如果夹具附件目标匹配成功,那么夹具附件将和操作号、夹具名称、夹持表面等一起加入夹具明细表中。

还有一类谓词用来评价所选用夹具是否合适,如谓词“ fix-chuck ”用于从夹紧误差方面来评价三爪卡盘的选用。这一谓词用来计算工件和工具的变形,从而估算出工件的径向误差和不同心度。这一谓词和其它谓词的不同之处在于,如果选用的夹具是谓词所指定的三爪卡盘,那么它的匹配总是成功的,并给出估算误差。

6 4 2 实例

图 6 .8 为一回转体零件的示意图,该零件由 8 个特征组成,可采用 Prolog 子句型式表示如下:

```
feature(1, fce , lft , 0,0,0,4,0,0,4, y [ +0 .023 , - 0], cir )
feature(2, hor , ext , 0,4,10,4,10,8,0, y ,[ +0 .01, - 0 .01], cir )
feature(3, ver , ext , 10,4,10,16,0,0,12, n ,[ ], cir )
feature(4, hor , ext , 10,16,18,16,8,32,0, n ,[ ], cir )
feature(5, ver , ext , 18,16,18,4,0,0,12, n ,[ ], cir )
feature(6, hor , ext , 18,4,28,4,10,8,0, y ,[ ], cir )
feature(7, fce , rgt , 28,4,28,0,0,0,4, n ,[ ], cir )
feature(8, hor , int , 0 .1 5,28,1 5,28,3,0, y ,[ ], cir )
```

式中的表面类型项“ fce ”表示端面、“ hor ”表示水平表面、“ ver ”表示垂直表面。表面形状项“ ext ”表示外表面、“ lft ”表示左面、“ rgt ”表示右面、“ int ”表示内表面。

第 1 个特征面子句的含意为:特征面代号为 1,左表面,起点坐标 $X=0$, 起点坐标 $Y=0$, 终点坐标 $X=0$, 终点坐标 $Y=4$, 表面长度 = 0, 表面直径 = 0, 表面高度 = 4, 需要加工, 与表面有关的公差为:上偏差 +0 .023、下偏差 0, 截面形状为圆形。

图 6 .8 一简单回转体零件的示意图

由 CAPP 模块,可得到零件的加工工艺,并用 Prolog 子句形式表示如下:

operation(10, face , blank ,[7],120,0 .01,0 .8)
operation(20, turn , rough ,[6],150,0 .15,1)
operation(30, turn , finish ,[6],200,0 .05,0 .1)
operation(40, turn , finish ,[5],180,0 .05,0 .2)
operation(50, turn , finish ,[4],180,0 .05,0 .3)
operation(60, turn , rough ,[2],150,0 .15,1)
operation(70, turn , finish ,[2],200,0 .05,0 .1)
operation(80, turn , finish ,[3],180,0 .05,0 .2)
operation(90, drill , blank ,[8],140,0 .005,0)
operation(100, ream , blank ,[8],200,0 .005,0)
operation(110, part , blank ,[9],160,0 .004,0)

其中，“ face ”代表车端面、“ turn ”代表车外圆、“ drill ”代表钻孔、“ ream ”代表铰孔、“ part ”代表切断。

有了上述加工工艺后,就可以利用本专家系统进行夹具的选择。在初选时,系统选出所有可能的结果,然后再通过优化程序进行优化处理。下面就是上述零件的初选结果:

Operator	Fixture name	Brand	Holding surface
10	three paw chuck	greatwall	2
10	three paw chuck	greatwall	4
20	three paw chuck	greatwall	2
20	three paw chuck	greatwall	4
20	turn- between- centers	xx	1,7
...			
110	collet	milwaukee	6
110	collet	acetools	4

在得到了所有可能的解以后,再利用优化方法为每一操作找出最合式的夹具,下面给出了最终的优化结果:

Operator	Fixture name	Brand	Holding surface
10	three paw chuck	greatwall	2
20	three paw chuck	greatwall	2
30	three paw chuck	greatwall	2
40	three paw chuck	greatwall	2
...			
100	collet	milwaukee	2
110	collet	milwaukee	4

6.4.3 系统的运行环境

本系统采用 Borland 公司的 Turbo Prolog 语言编写,可在微机上运行。由于程序中的有关文件,如工艺设计文件、夹具信息文件、特征文件和制造信息数据库文件等都是 ASCII 文本文件,它们可用任何通用的文本编辑器进行编辑。

6.5 CAPP 专家系统及其研究动向

CAPP 在我国的发展历史虽比国际上晚,约开始于 80 年代初。但发展相当迅速,现已形成了一支遍布全国的人数众多的研究开发队伍,其研究水平在许多方面与国际上相比并不落后。目前已开发的、投入运行的系统主要是基于成组技术的派生式系统或半创成式系统。对于 CAPP 专家系统,我国技术人员不仅很好地掌握了系统的开发原理与方法,而且在各行业工厂,针对各种不同的零件对象开发了好几十个 CAPP 系统,也不乏在实际中应用成功的。其中比较成功的系统主要是用于标准化或规范化程度较高的回转体零件和箱体类零件,因为这类零件的信息易于表达,知识的收集和表达也比较容易。

工艺设计由于其自身的特点,各行业之间,甚至同一行业的各企业之间都有很大差异。这主要是由于各企业都有自己长期以来形成的传统、经验和习惯。各工厂的设备条件对加工工艺也有很大影响,工艺人员的设计水平更是因人而异。因此,迄今为止,绝大部分 CAPP 系统都是针对各个厂家,甚至某一类零件开发的,其应用范围有很大局限性。随着市场经济的发展,多品种少批量的生产模式已成为机械制造业的主要趋势,传统的 CAPP 难以适应不断变化的加工对象和制造环境。

传统的 CAPP 系统的开发往往从最基本的模块开始,开发周期长,且大多数是非开放性系统,用户一般无法对其性能进行扩充和修改,难以二次开发。传统的 CAPP 系统的另一个主要问题是系统的柔性差,当产品发生变化,或者制造技术、设备环境发生变化时,系统难以适应。造成这种现象的原因,除了一些重要的理论问题和技术问题尚未很好解决以外,主要是传统的开发方法采用了针对特定的具体条件进行系统分析与设计的开发策略。这样,一方面从数据结构到控制逻辑都与具体问题关系密切,结果造成系统的开发者每次都要完成很多基础工作,降低了开发效率;另一方面,系统的使用者的自由度很小,不利于系统的完善与扩充。

随着工艺设计自动化理论、人工智能和计算机技术的发展,人们对 CAPP 系统的开发方法提出了更高的要求。为了提高效率,缩短开发周期,应避免过去系统开发过程中存在的大量低水平重复劳动,以提高开发起点。为此,国内一些高校和研究单位对此进行了不少工作,其目的在于使系统的数据、知识和控制三部分相互分离,建立一套比较完善的 CAPP 开发工具或 CAPP 开发环境,从而实现系统的开发、维护、扩充、移植的工具化。

6.5.1 以知识为基础的 CAPP 系统

CAPP 专家系统也称为以知识为基础的 CAPP 系统,一个 CAPP 专家系统必须具备

大量的事实和加工工艺规则。不仅如此,系统还必须具备一定的柔性,以便及时更新事实和规则,使系统可以不断更新,从而适应不断发展的生产环境。

在一般的创成式 CAPP 系统中,制造工艺知识被一行行地写入程序中。如果想要修改事实或规则,就必须修改程序。传统的 CAPP 由于难于修改而大大局限了其应用效果。

CAPP 专家系统和一般的创成式 CAPP 系统相比,有以下几个显著优点:

(1) 可括展性 以知识为基础的 CAPP 系统,由于带有可编辑的事实库和规则库,因此可以很方便地增加、删除和修改其中的事实和规则。在生产条件改变时,如更新设备、改进工艺、采用新技术时,专家系统可以通过修改知识库的办法来适应这些变化了的外部环境而不必修改源程序。

(2) 透明性 专家系统对用户有更友好的用户界面,系统可以随时向用户提供运行的过程和决策的理由,向用户作出各种解释,所以系统的工作是透明的。

(3) 可移植性 和一般的创成式 CAPP 系统相比,CAPP 专家系统可用修改事实库和知识库的办法来适应不同企业的需要,比较适合于推广应用。

(4) 学习功能 少数以知识为基础的 CAPP 系统还可以利用某种学习程序而具备一定的学习功能。

任何以知识为基础的系统的质量主要取决于其中的专家知识,把专家知识变为以知识为基础的系统可以使用的形式也称为知识获取。知识获取是一件费时、费力且易于出错的工作。但也是一项十分重要的工作。

6 5 2 CAPP 开发(工具)环境的研究

工艺设计由于其本身的特点,各行业之间,甚至同一行业各企业之间都有很大差异,这主要是由于各企业都有自己长期以来形成的传统、经验和习惯,各工厂的设备条件对加工工艺也有很大影响。工艺人员的设计水平和经验更是因人而异,因此,迄今为止,绝大部分 CAPP 系统都是针对各个厂家,甚至是某一类零件开发的,其应用范围有很大局限性。随着市场经济的发展,多品种、少批量的生产模式已成为机械制造业的主要趋势。传统的 CAPP 难以适应不断变化的加工对象和制造环境。

传统的 CAPP 系统的开发一般是从基本功能模块出发,起点低,软件的开发周期长,使用 CAPP 系统很难迅速发挥效益。

基于上述原因,为我国广大工矿企业研制一种通用的 CAPP 开发工具或开发环境已成为当前 CAPP 研究的热点之一。如基于 DBMS 的工艺设计的研究和采用可重用构件库等办法来实现的 CAPP 开发环境的研究等,从不同角度探讨了 CAPP 开发工具或开发环境的实现,为 CAPP 实现通用化、工具化研究提供了良好的借鉴。

6 5 3 动态 CAPP

目前 CAPP 专家系统的研究大多集中于独立的 CAPP 系统与 CAD 的集成。这种 CAPP 与生产计划相独立,不考虑车间生产的实际动态状况,其生成的工艺在车间实际生产环境中产生了一些实际问题,这些问题主要有:由于生产计划要在工艺设计之后完成,

而设备的负荷又不平衡,为此不得不改用替代设备以致改变工艺的情况时有发生。特别是在一些柔性制造系统中,由于工艺设计缺乏柔性,因此系统中某台设备的故障往往会造成整个系统的停产,从而大大降低了系统的柔性。

实现工艺设计与生产计划的集成,达到两者更好地协调,是解决工艺设计与实际生产脱节问题的必由之路。动态 CAPP 的特点是在进行工艺设计时,充分考虑车间资源的动态状况,这样制定的工艺设计结果,符合车间的生产实际,是解决工艺设计与生产计划之间不兼容问题的出路,也是目前 CAPP 研究的热点之一。

目前研究的动态工艺设计具有以下内涵:

(1) 考虑车间资源的静、动态情况,提供多个工艺方案。车间资源的静态情况主要是指车间环境中的设备种类、数量、性能以及操作人员的具体情况;车间资源的动态情况主要是指车间中设备的动态负荷,性能状况等。

(2) 实现工艺设计与生产计划的集成,能与计划共同作用产生优化方案。

在目前开展的关于动态工艺设计的研究中,有人提出了一种“伪实时工艺设计”的思想。在“伪实时工艺设计”中待设计的对象被分为三类:第一类还没有进入生产准备过程,对这类对象重新进行工艺设计和规划引起的浪费(工艺设计和规划的成本)在计算机环境下是很小的,因而允许进行多次设计;第二类已进入制造准备过程而未进入制造过程,这类对象不适宜重新进行工艺设计和规划,因为这可能引起较多的人力和物力的损失(如工装设备等),但允许进行车间实时规划;第三类已进入制造过程,处于生产控制之下,重新设计这类对象的工艺可能造成较大的损失。工艺设计被分为宏观工艺设计和微观工艺设计两个阶段。规划被分为预规划和最终规划两阶段。资源状况不仅包括静态状况和实时动态状况两种,还提出了预期动态状况概念。“伪实时工艺设计”可以按某时间周期进行,也可以按某设计对象运行,或在新增加若干设计对象时运行,或在生产其它环节反馈某些信息时进行。

“伪实时工艺设计”通过与车间规划相互作用,根据预期的资源未来的动态状况,对若干设计对象进行总体的优化设计,因此可以克服实时工艺设计的短见决策,削弱工艺设计与执行之间的时间差对工艺规程可用性的影响,解决因工艺设计与规划在任务上的差别所带来的问题。

6.5.4 面向并行工程的 CAPP

并行工程是 80 年代末由美国国防工业系统提出的一种新兴的制造概念和生产哲理。由于它在产品设计阶段就集中研制产品各环节的有关人员对产品性能和进程的设计,同时作出评估,及时改进设计,以得到大家满意的设计结果,同时对下游各技术环节作出准备和操作,达到各项技术设计和制造一次成功,因此既保证了产品的设计和制造质量,又能大大缩短新产品的研制和开发周期,增强企业的市场竞争能力。它已成为当前国内外制造业和学术界研究和实施的热门课题。

实施并行工程的通常的技术组织措施是建立各专业的协同工作组,即以 TeamWork 的方式开展工作。这种多专业并行协同小组要求将各类人员,包括设计和工艺人员组织在一起,但在具体生产环境下,要经常维持这种组织形式和工作方式是困难的。有关人员

不可能始终在时间和空间上集中在一起,他们还有各自独立的工作环境和任务,这就要求建立在计算机网络基础上的 CAD 和 CAPP 的集成。

工艺设计和产品设计同步进行是并行工程环境中 CAPP 与传统的串行设计方法中 CAPP 的主要不同点。采用智能化的 CAPP 系统,可以大大缩短工艺设计的时间,通过计算机网络,还可以实现 CAPP 和 CAD 之间的信息交换,把对产品设计的可加工性和制造成本等工艺性能反馈给设计系统,使工艺设计从原来的设计和制造的单向信息通道变为双向信息通道,以便及早修正设计中存在的工艺问题,使设计满足可制造性要求。对产品设计可加工性评价是并行工程的 CAPP 的主要特点。

在 CAD 进行产品设计的过程中,通过 CAPP 系统接口模块调用 CAPP 专家系统。CAPP 专家系统在工艺设计的同时对 CAD 设计结果进行分析,分析内容可依据 CAD 设计的不同阶段进行相应的分析与评价,结果反馈给 CAD 系统,确定设计的加工性能,以达到优化设计的结果。

面向并行工程的 CAPP 专家系统的另一特点,是它不仅和其上游的 CAD 保持密切的关系,且和其下游的生产计划、制造、装配、可靠性、可维护性等各功能模块之间保持密切的信息交换关系,以保证整体的最优化,这是面向并行工程的 CAPP 的另一主要特点。

第 7 章 制造过程的智能 监视、诊断与控制

7.1 概述

随着科学技术的发展和多品种小批量自动化生产的要求,制造过程的智能监视、诊断和控制已越来越受到人们的重视,并已在一些工厂中得到实际应用。由于现代制造过程是一种集多种高新技术为一体的现代化制造系统,为确保其可靠高效地运行,必须对其运行过程进行实时监视,以及时发现运行中的故障,并对其进行诊断和处理。监控系统质量的好坏,将直接关系到制造过程能否正常运行。因而,监控系统已成为现代制造过程中不可缺少的重要环节之一,也是世界范围内广泛研究的热门技术课题。在我国,有许多单位开展了这方面的研究工作。

当今是信息化时代,在制造过程现代化的过程中,已大量涌现以计算机为核心的监视和控制相结合的实用系统。伴随着这种系统的发展,一些先进技术,如信号传感技术、数据处理技术及计算机控制技术正在飞速发展。

电子测量仪器、自动化仪表、自动化测试系统、数据采集和控制系统在过去是分属于各学科和领域各自独立发展的。由于制造过程自动化的需求,它们在发展中相互靠近,功能相互覆盖,差异缩小,其综合的目的是为了提高人们对制造过程全面的监视和控制等多方面的能力。与此同时也对监视、控制技术本身提出了高技术的要求,如高灵敏度、高精度、高分辨率、高速响应、高可靠性、高稳定性及高度自动化智能化等。

现代的监视与控制系统,正在向智能化方向发展。所谓智能,是指能随外界条件的变化,具有确定正确行动的能力,也即具有人的思维能力及推理、作出决策的能力。而智能化的仪表或系统,可以在个别的部件上,也可以是局部或整体系统上使之具有智能的特征。例如,在以微机为核心的监视与控制系统中,可以实现最优控制、自适应控制等智能监视与控制功能,也可采用智能计算机和人工智能工程等,实现更为高级的智能化。

现代监视与控制任务更多地涉及到系统的特征,在系统内部,若要设立多台微机,则这些微机往往不是互不相干的,而是要构成相互联系的整体,这就形成了各种多微机的系统。即使是使用单台微机进行集中控制,也要通过标准总线和各个部件发生联络,由此发展了各种集散式、分布式的数据采集和监控,以适应系统开放、复杂工程及大系统的需要。在向系统化发展的同时,还需要涉及到系统部件接口的标准化、系列化与模块化,以便构成通用的整机。由上述可知,智能监视与控制系统是在数据采集处理系统与计算机控制系统的基础上发展综合而成的。有关系统的构成将在以后各节中逐步介绍。

7 2 接口、显示与控制器

7 2 1 智能传感器

传感器位于被测对象之中,在测试设备的前端位置,是构成系统的主要窗口,为系统提供赖以进行处理和决策控制所必须的原始信息。对于一个以微机为核心的系统来说,微机如人的大脑,而传感器则像人的“五官”,人要耳聪目明,系统也必须感觉灵敏,精确无误。传感器在监视系统中的位置见图 7.1 所示,是联系非电子部件与电子部件的桥梁,故它是实现制造过程的智能监视、诊断与控制的重要环节。

图 7.1 传感器在监控系统中的位置

当今信息的传输、信息处理与信息控制已相当发达并已通用化。某个具体过程、物态的动态监视或控制能否实现,归根结底为能否找到一些恰当的传感器可真实地、迅速地、全面地反映该物态或过程的特征,并把它变换成便于识别、传输、接收、处理和控制的信号。若是可以做到,那么一般来说能够实现智能监视和控制;反之,只能说人们暂时不能实现对它的识别与控制,有待于进一步开发出更合适的新型传感器后再作考虑。所以传感器是决定能否实现制造过程智能化的关键,是新产品、新设备开发的首要问题。

传感器是直接感受被测量的一次仪表,处于整个测量系统的最前方,在它之后的电子测量电路称为二次仪表。由于微机中引入信息处理、自校正等功能,二次仪表常以高保真度再现传感器的输出,但却无法增添新的检测信息或者消除传感器所带来的错误信息,所以获取信息的质与量往往一次性地由传感器的特性决定。或者说,传感器很大程度上影响和决定了系统的性能,因此在设计研制某一系统时,不仅首先要寻求到合适的传感器,而且要选择出更为优异的及成本低廉的传感器。此外,同一传感器由于使用方法不同,接口方法各异,调整补偿的措施不同,也会使传感器的性能指标发生很大变化。

传感器用来直接感知被测物理量,把它们转换成便于在通道间传输或处理的电信号。更明确地说,传感器应具有三方面的能力:一是要能感知被测量(大多数是非电量);二是变换,仅把被测量转换为电气参数,而同时存在的其它物理量的变化将不受影响或影响极微,即只转换被测参数;三是要能形成便于通道接收和传输的电信号。因而一个完整的传感器,应由敏感元件、变换元件和检测电路三部分构成。对于有源传感器,还需加上电源,其结构框图如图 7.2 所示。这种组成形式带有普遍性。但也不是所有的传感器结构都要由三个部件联合构成,在所谓直接变换的情况下,敏感元件和变换元件合为一体。如热敏

电阻,它可以直接感知温度并变换成相应电阻的变化,通过检测电路就可以产生电压信号输出。

图 7 2 传感器一般组成

从传感器的发展过程看,最初仅是测量热工和电工量,逐步发展到机械量、状态量、成分量、生物量,目前已发展到可以代替人的五官感觉,甚至可检测到人的五官感觉不到的微观量;并且又从单参数的检测发展到多个参数的扫描检测,从单维数据到二维图像、三维物体识别,甚至四维数据的识别。传感器的品质,也从廉价的简单传感器、结构型传感器、物性传感器发展到现今的智能传感器及整机一体化的集成传感器等。为了从五花八门的传感器中选择出最适用的传感器,对它们进行分类是必要的。

传感器的分类方法很多,从传感器的应用和使用的角度看,了解下面几种分类法是有益的。

1 . 按传感器检测参数分类

按要求检测的参数的类型来分类,这样涉及到的面很广,不可能全部列出,仅根据一些常用的过程参数归纳于表 7 .1 中。

表 7 1 被测参量传感器分类

参 量	传 感 器
几何尺寸	厚度传感器,CCD 图像传感器等
速度	转速传感器,角速度传感器,线速度传感器等
加速度	加速度传感器,振动传感器,角加速度传感器等
力	应变传感器,压力传感器,扭矩传感器,张力传感器等
流量	流量传感器,流速传感器,液位传感器,液压传感器等
化学量	成分传感器,PH 传感器,湿度、密度传感器等
光和放射性	等光传感器,光纤传感器,射线传感器
温度和热	热敏传感器,高温传感器,红外传感器等
磁	霍尔传感器,磁敏电阻,核磁共振等
气体、温度等	气敏传感器,温敏传感器等

2 . 按传感器方法分类

传感方法往往基于某种物理效应或材料的特性,使传感器完成能量变换,或者引起电

的参数变化,形成与被测量成比例的输出,可归纳为以下几种类型:

(1) 能量变换方式 能量是从被测系统提取,转换为一种与它等价的电的形式(中间也有能量的损失)。例如电磁感应式压电晶体、热电偶、光电池等,一般无需外加电源。

(2) 阻抗控制方式 由被测物理量变化引起相应的电路参数的变化(如电阻变化、电容变化、电感的变化等),从而可以通过检测电路形成电流和电压变化的输出。这类传感器如热敏电阻、湿敏电阻、光敏电阻等,这类方式要外加电源才能形成检测的电信号输出。

(3) 平衡反馈方式 具有反馈的特性,而这种反馈特性是输入物理量和一个与它对抗的电量相平衡的效应,指示出达到平衡所需的势值就给出了被测物理量的值,平衡反馈式要应用更复杂的效应。

3. 按传感器输出电信号形式分类

传感器的输入量大多数是随时间作连续变化的物理量,其输出量也以模拟电量形式居多。近年来由于数字技术发展,以微机为基础的系统增多,从提高检测精度着眼,直接数字式传感器、频率变化式传感器、脉冲参数变换式传感器也有发展。因而从输出电信号形式又可分为以下几类:

(1) 开关式传感器 开关式传感器工作特性为:当输入物理量高于某一阈值时,传感器处于接通状态,以低电平(或高电平)输出;当输入物理量低于某一阈值时,传感器工作在另一种开断状态,输出高电平(或低电平),所以输出是以高、低电平形式变化,例如限位开关传感器。

(2) 数字式传感器 它又可分为直接数字传感器、频率式传感器和脉冲传感器。直接数字传感器输出经过编码的数字量,例如光电编码盘;频率式传感器输出的信号反映频率变化,可以直接用数字频率计来测量;脉冲式传感器,输出信息反映在脉冲的个数或参数的变化上。数字式传感器共同的特点是精度较高且便于和微机接口连接。

(3) 模拟式传感器 传感器输出以各种连续量的形式出现,可以是电压、电流、电阻、电容、电感等。连续变化量与数字系统连接,需通过模拟通道的转接。

传感器还可按物性材料或其它方式分类。

近年来传感技术的发展非常迅速,从结构型传感器发展到材料型物性传感器,又发展到带有信息处理的智能型传感器及整机化集成传感器,各种新型传感器层出不穷。从国内外发展的总过程看,传感技术有两种发展趋势:一方面是大力发展各式各样的廉价传感器以适应消费领域,如家用电器、汽车运输等的需要;另一方面则是面向可靠性、抗干扰能力强、高精度、高速度,并且有大量附加功能的第一流传感器方向发展。在这一方面已出现了多功能一体化的传感器系统,发展了以完善传感器为中心的广义传感器,而集成传感器和智能传感器就是发展中的一个重要侧面。

智能传感器是一种带有微处理器的敏感探头,它是兼有信息检测和信息处理功能的传感器。它以集成化为基础,将敏感检测、信息处理及微机集成在一块芯片上,或者装在一个外壳里,所以从制造技术而言,是一种采用高技术的硅集成电路传感器。有的智能传感器甚至还包括了参数的调整过程控制和过程优化等更加复杂的检测控制系统。智能传

感器的智能作用,可归纳为以下几个方面:

(1) 提高传感器的性能 通过微机信息处理和集成工艺技术,可以实现自动校正和补偿。例如对传感器的线性度、重复性、分散性及老化效应自动校正;对集成于一片的多个传感元件零位自动补偿;自动选择合适滤波参数,消除干扰与噪声;自动计算期望值、平均值和相关值;还可以根据传感器模型作动态的校正,以实现高精度、高速度、高灵敏度宽范围的检测,甚至可实现无差测量。

(2) 自检与自诊断 自检是通过合适的测试信号或监测程序来确定传感器是否完成自身的任务。自诊断即在传感器损坏前,在作正常测量的间隔,通过检测一些特征量,并将其与被保护存储起来的状态、参数及期望值进行对比,以判断与分析其是否接近损坏。可按照一定顺序对关键部件进行检测,以确保测量的高可靠性。

(3) 多功能化 智能传感器为了对工况作出优化处理,需要同时检测多个量以便作出相关分析与处理。因此将多个传感元件甚至相应的检测与信息处理电路都集成在一块芯片上,使功能大大超过仅检测某一个物理量的普通传感器。因此,智能传感器扩大功能的一个方面是形成多元传感系统。目前已有产品能做到在一块智能传感器芯片上检测 8 个参数,例如某一对气体样品进行分析的智能传感器,其中集成化多元传感器阵列如图 7.3 所示。该传感器采用自动分时顺序扫描方式检测各种气体成分 $x_1 \sim x_i$ 的浓度值 $A_1 \sim A_i$ 并存储。该系统考虑到每一个半导体气敏传感元件对所有的气体成分都有不同程度的敏感,这样使用 n 个不同的气敏元件组成传感器阵列 a_{in} ,就可以得到 n 个传感元件的信号谱 $S_1 \sim S_n$,每个信号谱中包括 i 条信息。最后由微机进行模式识别,从一系列信号谱中得到气体成分的浓度答案 $A_1 \sim A_i$ 。这样测得的 $A_1 \sim A_i$ 值较使用多个分立的传感器测得的值更接近实际值,从而可消除各种检测元件存在的相互影响,并充分发挥微机分析综合的智能作用。

图 7.3 集成化多元传感器阵列

扩大功能的另一方面是使智能传感器可对多个测量值作静动态处理、运算,进而实现简单的调节与控制算法。使各智能传感器成为一个测量的前端机,能直接与上层机(或协调主机)通过数据通讯联系,实现集散式分级控制。

综上所述,较完善的智能传感器实际上也就是目前研究的智能检测与控制系统。只不过它不是分散的部件,而是集成于一芯片或联系在一个管壳内的统一整体,是高度集成化的产品。这样把各个环节集中于同一基片上,不但使传感器处于相同温度下,有利于作温度补偿或修正,而且必将大大便于用户,节省新产品的开发时间、调试校验时间,促进系统向更高级的智能化、网络化、小型化发展。

图 7.4 所示为美国霍尼维尔公司推出的 DSTJ3000 差压静压、温度三维测试和调整系统的智能传感器。该传感器内有差压、静压、温度三类敏感元件并集成在同一 *N* 型硅片上,还带有多路转换、传感脉冲宽度调制器、微处理器和模拟输出及数字输出等部件,并具有对信号远距离传输和调整的能力,也具备智能传感器一些基本的智能特征,如自诊断、自补偿、自校准等。

图 7.4 DSTJ-3000 智能传感器框图

7.2.2 信号变换与接口电路

在制造过程智能监视系统中,在传感器到微机之间,尚需要恰当的信号和接口电路,微机才能定时地采集数据到机内进行处理。信号变换是指保持原有的信息不丢失,而以不同的信号形式或信号的不同特征来重现。在对信号往往施行多次变换或反变换,但都具有一定的目的。传感器可看作信号的第一次变换,是把非电物理量变为电量。后续信号变换是为了便于微机接口、数字信息处理或信息的远距离传送等目的设计的。

信号变换级与微机的接口电路,要根据所选择的传感器类型、传感器和微机检测中心之间的距离以及整个系统性能指标的要求来设计选定。大多数工业应用中,传感器输出多为连续的模拟量电信号,需通过离散化处理,将其变换为便于微机处理的数字信号。在离散变换之前,有时还要作其它形式的预变换,如将电流或电荷量变为电压量,适当放大并滤去噪声干扰等。变阻抗类传感器输出信号往往是交流信号。若信号幅度反映了信息量,可以通过检波或整流变换电路,将其转换为直流信号再作离散处理,对低频交流信号也可以直接作离散化处理;若信息是反映在交流信号的频率变化中,则可将其整形为脉冲

信号,使用数字式频率计的计测方法处理;若传感器本身为数字式传感器,可输出开关量脉冲信号或已编码的数字信号,则仅需进行脉冲整形、电平匹配或数码变换就可以和微机接口。

当传感器和微机检测系统之间距离较近,使用屏蔽电缆或双绞线就可将传感信号与信号变换检测电路相联。若距离较远,或者还存在一些不能直接相连的其它因素,则往往要将传感信号调制成高频载波,通过信道或其它媒介传输到计测中心。

图 7.5 所示为一个单传感器输入情况。传感器通过屏蔽电缆与监视系统相连,或经预处理(虚线框所示)之后与监视系统相连。其中预滤波环节根据系统存在干扰及信号频谱的情况,作频带宽度的压缩,以便于采样频率的选取。系统中的放大器用于传感器到 A/D 转换之间的电平转换。滤波器处于放大器之前,可以更好地抑制噪声,提高信噪比。

图 7.5 单模拟输入采集通道框图

实际的智能监视系统常要求同时完成多个物理参数的检测,即所谓巡回检测。在巡回检测时,分时逐点地完成检测。使用多路模拟开关,定时扫描接通各检测点,形成多路模拟输入分路器。多路模拟开关可以插到各路信号预处理与采样之间,以共享采样保持、ADC 及微机监视系统。如图 7.6 所示。

图 7.6 一种多数和数据采集通道框图

有时也可根据被处理信号的不同情况,插到数据采集通道的其它各个不同部位,如放大之前(此时放大采用可变增益的程控放大器),采样保持之前或 ADC 之前均可。多路模拟开关处在系统比较前面的部位,则各路信号可以共享的检测电路较多,因此比较经济。但有时从提高某些质量指标出发,或各路电平差别较大无法实现共享,则放在通道的后方也是可以的。

多路模拟转换集成电路芯片,可以转接的路数是有限的,通常是 8 路。若需要转接更多的路数,可以采用多片。先由微机进行片选,而后再片内选址,选择某一路接通后面的

检测电路。多片模拟多路转换器的连接方法,尚可以采用多级分组的连接方法,以便扩展路数,如图 7.7 所示。

图 7.7 分组多路转换

采集的路数受限于数据采集的实时性、高速性以及各路信号之间交叉干扰等问题,在有些场合不得不采取多通道的接口方法。如图 7.8 所示。但多个采集通道的造价是昂贵的,因为 ADC 芯片通常较其它芯片价格高。

图 7.8 多个模拟通道

包含所有环节在内的模板结构的数据采集系统虽已有产品出售,但尚难以满足各式各样采集系统的全面要求。对于某些实际应用还经常要根据具体情况设计,比较理想而造价又不太高的数据采集系统,特别是当追求高质量系统或某些指标有独特的要求时,更要精心设计系统的各个环节。

在采集子系统的前端,需设置硬件预选滤波器,其作用有两个:其一是为了在前置的放大之前,预先抑制部分信号噪声干扰,以免和信号一起放大,可以有效地提高信噪比;其二是压缩频带,限制到小于 $0 \sim f_s/2$ (f_s 为采样频率)的范围内。由于信号和干扰噪声在时域上是混杂在一起的,但是在频域上两者具有各自不同的统计特性,所以一般来说,有用信号的频谱总是有限的,而噪声频谱却宽得多,甚至在高频段仍有很大值。

实现频域滤波的方法可分为两大类:模拟滤波方法和数字滤波方法。但由于预选滤波是置于系统的最前方,信号尚未被数字化,故多采用模拟滤波器。对于大多数工业与生活用的传感器,信息变化是很慢的,典型的仅 10Hz 左右,而干扰噪声频谱则分布在比较高的频区,故应采用低通滤波器,低通滤波器的设计可参阅相关书籍。

在多路监视系统中,需要由多路转换器转换各路模拟信号。多路模拟转换器(MUX)的集成芯片使用已很广泛,它是多路光触点双向 CMOS 模拟开关。一般来说,它相对于过去通用的干簧管触点开关,具有寿命长、速度高、便于程控及长期连续工作不会出现粘连问题等优点。

MUX 与多路模拟电压连接时,需要考虑是单端式或差动式连接。对于单端式连接可参见图 7 .9,在图中使用一片 AD7503 芯片,它可转接 8 路模拟输入电压;对于差动对称输入连接,可参见图 7 .10,在图中使用 3 片 AD7502,它是双路四通道的多路开关,可有四路输入。

图 7 .9 多路转换器的连接

图 7 .10 多路转换器的差动连接

MUX 的多路模拟信号切换是由提供给芯片的开关地址码(如 AD7503 有 8 个输入通道,它有三根开关地址码引脚 A_0 , A_1 , A_2 ,可通过外部软件或硬件的方法,顺序地提供对应的通道地址码)进行切换并在使能端 EN 为 1(AD7501 芯片使能端为低电平选中)时接通所需通道,图 7 .11 为采用两片 AD7503 扩展构成 16 路通道的多路模拟开关与微机

的连接(属于软件切换)。

两片 AD7503 的 EN 端由 74LS139 译码器产生的片选信号 \overline{CS} 及 D_3 位数据点线共同选择。当 $\overline{CS} = 0, D_3 = 0$ 时,选中上片 (AD7503),它的 3 位地址码通过数据总线送至 74LS175(4D 触发器)转送过来。若 $D_2, D_1, D_0 = 000$,则选通 U_0 输入通道,若 $D_2, D_1, D_0 = 001B$,则选通 U_1 输入通道,以此类推分别选中 $U_0 \sim U_7$ 8 个通道;当 $\overline{CS} = 0, D_3 = 1$ 时,则选中下片 (AD7503),可由数据线上 $D_2 \sim D_0$ 3 位代码,分别选中 $U_7 \sim U_{15}$ 8 个通道。

图 7 .11 软件控制切换的 16 路 MUX 与系统连接图

软件切换可做到灵活多样的控制,但软件切换要占用 CPU 的工作周期,并且切换速度较慢,为此在高速多路采集系统中,常用硬件计数器自动提供顺序的切换地址代码,图 7 .12 示出了这种连接方法。使用一块 16 进制硬件计数器,输出 4 位代码,其初始状态清零,之后 ADC 每转换一次数据,使计数器增 1,顺序地产生地址代码以切换 16 路模拟信号。

图 7 .12 硬件控制切换 16 路 MUX 系统连接图

采样/保持环节通常处在数据放大(或多路转换器)与 A/D 转换之间,作为逐次比较式 A/D 转换器的前端,用以对快速变化信号设立摄取采样瞬时窗孔,以实现高精度的转换,因此它对系统的精度也有一定影响。典型的采样/保持电路如图 7.13 所示。逻辑控制由电压 U_c 控制模拟开关 S 的闭合,控制电路运行处于采样或保持两种模式。在采样模式中,S 闭合,输出跟随输入变化到采样值;在保持模式中,S 断开,输出将保持在保持命令发出瞬时的输入电压值,直到数字控制输入端输入了下一次采样命令为止。理想情况是,在采样期间输出完全跟随输入变化,而保持期间输出保持恒定不跌落。但实际上由于电路的瞬态响应及电路元器件的不理想,采样/保持过程瞬态变化是很复杂的。

图 7.13 典型的采样/保持电路

在使用采样/保持芯片时,重要的是选择适合系统的芯片,采样/保持集成芯片可分为三类,应根据系统要求进行选择。

通用采样/保持芯片,如 AD582,AD583,LF198。

高速型采样/保持芯片,如 HTS-0025,HTC-0300,SHA-2A 等。

高分辨率型采样/保持芯片,如 SHA1144,SHA-6 等。

通用型的采样/保持芯片,如 AD582,它是一个由高性能的运算放大器、低漏阻的模拟开关和一个由片型场效应管集成的放大器组成。全部电路集成在一个芯片上,只有保持电容器是外接的。AD582 的主要参数见表 7.2。为了能对比三类采样/保持芯片,表中也列出了高速型 HTS-0025 和高分辨率型的 SHA1144 的主要参数,可以看出高速型 S/H 芯片的特点是捕捉时间和孔径时间相对其余两者非常小,为 ns 级;而高分辨率采样/保持芯片的特点是孔径时间不定性和保持电压下降率特别小。

表 7.2 三种采样/保持芯片参数对比

指标型号	AD582 (通用)	HTS-0025(高速)	SHA1144 (高分辨率)
输入电压	差动 30 V	+/- 2 V	+/- 10 V
输入旁路电流	3μA	15nA	0.5nA
输入电阻	30 M	10^{10}	10^{11}
捕捉时间	($C_H = 100\text{pF}$), 6ms ($C_H = 1000\text{pF}$), 25μs	20ns	5μs
孔径时间	150ns	6ns	50ns
孔径时间不定性	15ns	20ps	0.5ns
下降率	下降电流 1nA	0.2mV/μs	1μV/μs
输出电阻	5	3	<1

数据采集系统中使用的 A/D 转换器,大部分为已集成化的芯片,故无需用户再进行单独设计。然而根据系统的需求选择恰当的芯片以满足系统性能指标的要求,并正确实施使用仍是必须的。

A/D 转换器性能指标有多项,现仅根据在微机采集通道中的使用,选择必要的几项进行分析。

(1) 输入与输出规格 ADC 芯片的输入规格,指明其转换范围,通常市场上的 A/D 转换器有几种标准规格。单极性的,如 0V ~ 5V, 0V ~ 10V; 双极性的,如以 0V 为中心的 ±5V 或以 0V 为中心的 ±10V 等,输入端信号应落在它的满量程范围之内。

与输入量相对应的是输出数字代码规格,也分为单极性代码与双极性代码。单极性代码的数字输出信息中不包括符号信息,所以比较简单,可以是纯二进制代码或 BCD 码。以显示功能为主时,采用 BCD 码比较方便,若需要直接和计算机连接,则用二进制代码比较方便。如果是按双极性方式工作,则在数字输出信息中包含符号信息,所以复杂些。

(2) 分辨率 分辨率决定了 ADC 传输特性中可识别的量化区间的大小,换言之,它表示了对模拟量输入信号进行数字化能够达到的精细程度。一旦输入模拟信号的动态范围和要求分辨的最小电压已知,就可以对应地计算出要求 ADC 的最大输出数值和最小输出数值(除零以外)之比,也即

$$\frac{U_{imax}}{U_{imin}} = 2^n - 1$$
$$N = \log_2 \left(1 + \frac{U_{imax}}{U_{imin}} \right)$$

式中, N 为 ADC 芯片字长的位数。它反映了信号的分辨率,或者说分辨率可以由位数 N 间接表示,即

$$\text{分辨率} = \frac{1}{2^N}$$

(3) 精度 ADC 的绝对精度,是指对应于一个给定数字量的实际模拟量输入与理论模拟量输入之差,引起 ADC 误差的原因是多方面的,如增益误差、非线性误差和温度误差。

(4) 转换速度 转换速度是 ADC 将模拟量输入信号转换成数字信息的速度,是转换时间的倒数,对于不同结构原理的 ADC 芯片,其转换速度有很大的差别,可以从几十纳秒到几百毫秒,所以转换速度是选择 ADC 的重要指标之一。

(5) 电源及参考电源稳定度 ADC 芯片的参考电压及电源电压波动,相当于引入干扰。参考电压对芯片的精度有更为直接的影响,ADC 芯片内部所带的参考电源均已有稳压措施,故外接电源时应根据其精度要求加入相应的稳压措施。

(6) 接口情况 有的 ADC 芯片带有外三态控制的输出缓冲器,可直接挂在总线上;也有的不能直接与总线联接,而要通过其它接口芯片转接。数据输出形式有并行与串行之分,使用时应加以注意。

不同类型的 ADC,其性能指标差异比较大,且各有特色,故必须有一个较全面的了解,方能正确选择使用。以下仅从常用于数据采集及控制场合的 ADC 加以分析及比较。

1. 逐项比较式 ADC

这类芯片是目前数据采集系统中应用最为广泛的一种,它易于获得较高的转换速度,高分辨率及较高的精度,也易于和微机接口。其价格由精度和速度而定,相差很大。

目前各种速度和精度的集成化逐项比较式芯片已很容易买到。现列举一些常用的芯片如下:

(1) 8 位分辨率芯片 常用的 8 位分辨率芯片有 ADC0809, AD7574。它们内部都带有外控三态输出缓冲器,可直接与微机接口;也可经 I/O 口再与微机连接,并有内部时钟。一般价格低廉,使用方便。ADC0809 还可接收 8 位模拟输入。

(2) 10 位分辨率芯片 如 AD7570 和 AD7571。它们都带有外控三态逻辑。10 位数据可分别用两个字节提供 8 位数据总线,使用两个使能端,一个控制端(HBEN)使两个最高位有效;另一个控制端(LBEN)控制低 8 位输出。

(3) 12 位分辨率芯片 如 AD572, AD574 及 AD578,它们也都是由内部控制的三态逻辑输出,但三者情况不完全相同。

逐次逼近式 ADC 芯片使用时,常需配置相应的采样/保持电路。

2. 双积分式 ADC

双积分式 ADC 以能获得高精度、高分辨率和低转换速度为其特点,在一些有强电干扰,而信号变化极其缓慢的工业应用场合有较大应用价值。

单片双积分类型的 ADC 集成芯片产品虽不如逐次比较型普及与多样化,但也有商品出售,如 13 位的 CMOS 产品 AD7550, AD7552 等。

3. V/F 或 A/D 转换器(VFC)

VFC 是把模拟电压信号变换为可计测的数字频率器件。这种类型的 ADC 主要特点是:允许输入信号的动态范围大;转换时抗干扰性能强,具有转换的单调性;输出是串行频率信号,便于远距离传送和隔离。其缺点是转换速度很低,一般为毫秒级。目前 V/F 式转换器在微机与传感器接口的外围电路中应用较多。

单片的集成块 V/F 变换芯片(未包含频率计数部分)的产品也很多。目前一些公司的电压-频率变换芯片已做成 VFV 型。VFV 芯片,在外部加少许元件可构成 V/F 的变换;也可以把同一芯片改接少量的外部引线,构成 F/V 的变换,因而称之为 VFV 型。这类芯片的型号有 VFV9400, ADVF32 和 AD537 等。

VFC 型 ADC,由于工作速度很慢,转换前无需采样/保持电路。

4. 并行比较型 ADC

在前述三种类型的 ADC 中,以逐次比较型的转换速度最快,其转换时间通常可达几微秒到几十微秒之间,但速度再高就难以胜任。然而对一些瞬态特性的测试,往往要求更高的采样频率,此时必须尽量缩短 ADC 转换时间。并行比较型 ADC 就是一种快速 ADC,其转换速度可以达到几纳秒至几百纳秒,但要实现高分辨率就比较困难。目前已

有的集成芯片有 6 位的 ADC5010/AD6020,其主要技术指标为: 6 位分辨率,100MHz 转换速率。并行转换 ADC,需要配用高速的采样/保持电路,以缩短采样时间,否则系统的速度仍然提不高。

5. 串-并式比较型 ADC

为获得高速转换又要求有很高分辨率的场合,采用串并行比较 ADC,可以综合串式比较与并式比较两者的优点。串并式 A/D 转换也要配合高速的采样/保持电路。

ADC 芯片与微机系统接口中还有一些共性问题,如芯片数据输出接口问题、芯片转换时序控制、数据传送方式与采样/保持电路的连接以及接地连接等问题。

(1) 芯片数据输出接口 在完成 ADC 芯片与数字系统连接时,必须了解芯片内数据输出硬件结构情况,以便正确连接。现把常用 ADC 芯片数据输出情况,归纳为以下三种情况分别讨论之。

第一种情况,芯片数据输出端带有三态逻辑控制的缓冲器,并在芯片处有三态控制端引出。例如 ADC0809 芯片的“OE”控制端,此种情况下,它可以直接挂于微机的总线上。在 A/D 转换期间三态门未打开,呈现高阻悬浮状态,当 A/D 转换完毕,用微机的 \overline{RD} 线控制打开三态门,使被转换的数据送上数据总线。

第二种情况,ADC 芯片输出端没有带三态缓冲输出结构或虽带有三态缓冲器,但三态门的控制端是由芯片内部的时序线控制,即外部无法控制,如果此时时序不能与微机配合,就不能与微机总线直接相连,故必须通过外设 I/O 端口转接。这类芯片如 AD370, AD571 及 AD572 等。

第三种情况,ADC 芯片内带有三态逻辑电路,对外没有控制端,仍由内部逻辑控制电路控制数据输出时间,但其内部控制的时序能和微机数据总线的时序配合,无需外部接口电路就可以直接和微机总线相连,如 12 位分辨率的 AD574 等。

总之,A/D 变换器芯片的三态输出结构是与微机总线连接的必要条件,但不是惟一条件,尚需注意是否能配合微机读写时序,以及微机的输入方式。如查询或中断方式时,尚须状态口及中断向量入口配合。

(2) ADC 芯片与微机接口的时序配合 各种 ADC 芯片与微机接口的时序配合问题,主要归结如下:

- 由微机发出芯片所要求的启动转换信号;
- 由微机给出 ADC 芯片与总线是否连通的片选信号或地址有效信号;
- ADC 芯片转换结束信号,可作为微机输入数据前查询状态,也可以作为中断输入的请求信号;
- 由微机发出读信号以打开三态输出的控制门,使数据送上数据总线,完成一次数据采集输入;
- 转换时钟信号。

提供芯片上述几种信号的方法,并不局限于一种,但是都应满足 ADC 芯片正常工作的时序要求。

(3) ADC 数据输入方式 ADC 芯片转换所需的时间各不相同,往往要通过转换结

束信号给予指示。微机只能在它转换结束后,采用适当的数据输入方式,将数据读入,主要有延时等待、查询、中断及 DMA 方式。

等待方式:根据 ADC 既定的转换时间,可在对其发出启动命令之后,插入足够数量的“空操作”指令,以保证消磨掉转换时间之后,用指令完成输入操作。此时芯片的“转换结束”信号可以不予以理睬。这对于转换速度很快,消耗的时间不多时可以采用。

程序查询方式:此时应将 ADC 芯片“转换结束”或“数据准备好”信号送入至 I/O 端口的一位,用程序查询的方法来检查这一位的状态。当这一位状态变化时,读入数据。

中断方式:将“转换结束”或“数据准备好”信号作为中断请求信号,由中断接口管理排队,当微机响应此中断请求之后,转入中断服务程序中,再执行输入操作。还可以采用外部定时器定时的中断方式,将数据定时的采入,只要定时间隔大于 ADC 芯片的转换时间,则无需理睬“转换结束”状态信号。中断方式可以免去 CPU 等待及查询所消耗的时间。

DMA 方式:主要用于芯片转换速度高于 CPU 数据传送速度的场合。即采用 DMA 方式时,可以不通过 CPU,直接在 ADC 和微机的 RAM 之间进行数据传送。

(4) ADC 接地与去耦的连接 ADC 外部接地与去耦的不当,也会降低模拟信号的精度,ADC 芯片对外均有模拟地 AGND(或模拟公共点)和数字地 DGND,正确的接法如图 7.14 所示,它们分别作为外部模拟地和数字地的公共端,两者之间再短接,也就是以 ADC 的 AGND 作为模拟参考点,所有的电源去耦电容及模拟电路的地 GND 都应经过分开的印刷板的铜箔或导线联到这一点来。而数字地与逻辑电路的地在另一点相连。这样可免去地线上的干扰。此外,为了获得好的去耦效果,在每块芯片的 $\pm 15V$ 、 $+5V$ 的电源进线处必须外加电容器旁路。例如,可采用 $1\mu F$ 以上的电容,且这些电容安排尽量靠近器件。虽然在芯片内部每个电源进线端内,大多有旁路电容,但一般所接电容值较小,故外部再接一个较大的电容,就可以起到良好的高低频干扰的去耦作用。

图 7.14 正确的接地去耦连接

7 2 3 显示与记录设备

对于数据采集、处理或测控系统,其数据输出与记录工具在系统中也占据相当重要的位置。它涉及被检测的信息能否重现输出的中间或最后结果并以清晰、精确、可靠的形式表达和保存起来。

显示是指信息在检测预处理之后,立即以适当的形式表现出来。但是信息显示之后也就消失,不再重现。当信息需要长期保存、复现、后备处理应用时,就需要有相应的记录装置,信息显示或记录都希望直观明了。有些需要用字符、数字标明;有些则采用图表、图形、曲线并给以色差表明。故需根据信息的特征、数据量的多少及精确程度等方面适当选择显示与记录装置。

常用的显示设备有各类字符管、灯示管、CRT、存储式示波器、图示仪及模拟式或数字式万用表等。常用的记录装置,按记录介质的不同又分为硬拷贝和软拷贝,所谓硬拷贝就是以能用人眼直接观察、直接理解的信息形式记录下来的,如 X-Y 记录仪、点阵打印机等。而软拷贝的特点是以间接的方式,不能直接理解的形式记录存储信息,必须进行某种变换处理后才能直读。常用的软拷贝即是软磁盘、硬盘、磁带等,大都是以数字数据形式保存。

许多显示、记录装置,往往自成系统并且本身也是智能化的,带有微机反馈控制和多种功能集中在一起。当它用于生产的监视控制系统中,则只是系统的一个部件。它与主机的联系可以通过并行或串行接口方式,也可以采用总线通信的方式与主机交换信号。不论哪种方式都要求把数据信息转化成能显示、记录的特定信息形式。如二进制代码信息要能在 LED 数码管显示出对应的十进制数字,必须作二进制代码到 BCD 代码的转换,再把 BCD 码经译码器转换成段显示代码和位显示代码加于各位数码管上才能正确显示。以下仅介绍 CRT 显示仪的控制原理及接口。

CRT 显示图形的方法有光栅扫描的点阵式图像法和随机扫描矢量图像法。点阵式图像的概念是把图像看成点的集合,或者说是由像素组成画面,每个像素代表一位二进制信息或者 P 位二进制信息(指彩色显示或有深浅的灰度显示)。当 $P = 1$,即黑白显示时,一幅画面则由 $M \times N$ 个像素点构成, M 指纵向的像素点数; N 是指横向的像素点数。把一幅图像的所有像素一行一行顺序地存储到存储器 RAM 中,就在 RAM 中构成了一幅虚拟画面。当 $P = 1$ 时,虚拟画面所需存储容量的位数 $S = M \times N \times P$ 是一个字节 8 位的整数倍。若 M 与 N 比较大,则代表水平方向与垂直方向的可分辨率比较高,但虚拟画面所需存储器的容量也比较大。显然,彩色虚拟画面的存储容量大于单色显示的存储容量。

显示一幅画面,必须先由计算机将一幅画面的点阵图像信息先存储到 RAM 之中,显示时再由 CRT 控制把点阵图像信息逐点顺序地与水平、垂直同步信号复合成视频信号去调制显像管的亮度。当显像管电子束从屏幕左上角开始向右一行一行扫描时(由行同步信号控制),每个像素点都受点阵图像视频信号的控制,显示明暗(或不同彩色)的像点,待一幅画面显示完毕,电子束扫到屏幕的右下角,即逐行扫完一场后,在垂直同步信号控制下进行垂直回扫,再返回到第一行的起始端。

根据上述显示原理,CRT 的监视器与微机接口时,将考虑以下的接口环节,如图 7.15 所示。图中多路开关 SW 先打向 A 侧,使虚拟画面存储器 RAM 先与主机的总线相连,先将点阵图像的二进制信息写入 RAM。写完一幅画面后,再将多路开关打向 B 侧,使虚拟画面存储器与 CRT 控制芯片相连,由此芯片即可产生点阵图像信号、水平同步信号、垂直同步信号等混合视频信号,并将其连到 CRT 显像管的控制栅极以调制辉度。

所谓矢量图像法是把图像看作为矢量段的集合,由于任一条曲线都可以用许多首尾

图 7 .15 图形图像的显示记录接口

相接的短矢量来逼近,因此,可以在屏幕上先给一系列坐标点,然后控制电子束依次扫描,逐步产生首尾相接的短矢量,构成近似曲线。当矢量段足够短时,看起来就是很光滑的曲线。矢量要由矢量形成器产生。一个矢量可以用起点坐标和终点坐标来表示或者用矢量的起点坐标、矢量的增量体(X , Y)和增量符号表示,因此,只要给矢量形成器这些参数,就可以产生相应的矢量。在某些智能 CRT 中,带有丰富的图形软件,其中也包括把矢量图像信息转换成点阵式图像信息的软件,这样即使在光栅扫描的点阵式显示仪中也可以用矢量绘图。

矢量图像与点阵图像相比,矢量图像法画线速度快,画线质量好,形成的图像清晰,能实时修改图形内容,需要的帧存储器容量小,但是不宜于显示复杂图形;而点阵式图像可以显示各种复杂的图形包括立体图形。

CRT 显示西文字符(指英文字母、数字及符号等)的原理,也是点阵式显示法。一字符用 5×7 点矩阵或 7×9 点矩阵表示,如果显示比较复杂的字符,如汉字,要用 16×16 点阵或 24×24 点阵表示。但是由于计算机送出的字符数字信号均是 ASCII 码表示的,不是直接的点阵显示信息,所以还必须把 ASCII 码转换为点阵字符显示信息才能显示。这样,显示原理与接口电路与图形显示就稍有差别,因而显示设备常有字符显示模式(即文件工作方式)和图形显示模式之分,其分辨率要求也不一致。

要把 ASCII 码转换为点阵字符显示信息主要借助于字符发生器,字符发生器已有专用的商品。以 5×7 点阵表示的英文字母、数字、符号的字符发生器为例,其字符显示码存放在约 2K 字节的 ROM 中。CRT 在显示字符时,在刷新存储器内也是存放一帧画面的全部信息。通常一帧画面可显示 25 行字符,每行可容纳 80 个字符。这些字符行的位置、列的位置,均由 CRT 控制芯片控制。根据上述 CRT 显示字符的原理可以画出其控制原理框图,如图 7 .16 所示。通常 CRT 内部均带有自身的微处理器,这样 CRT 与主机的通信就可采用串行通信方式。

在工业检测系统中已愈来愈多地需要加入汉字信息,以增加信息的直观性,便于操作和普及应用。用 CRT 显示汉字信息也很方便,因为汉字也是一种字符,也可以看成是点阵构成的,所以可基本上仿照字符显示的方法处理。更详细的内容可参考有关参考书籍,本书不再详述。

最后附带说明,点阵式打印机作为硬拷贝记录设备时,其记录图形、字符及汉字原理与 CRT 显示图形字符及汉字的控制原理基本相同,图 7 .17 为打印字符控制接口电路。

图 7 .16 智能 CRT 内部结构与主机接口

图 7 .17 点阵打印机记录字符控制接口

7 2 4 可编程控制器

1 . 概述

可编程控制器(programmable controller)通常简称为 PC。为了和个人计算机(personal computer)相区别,有时也称为可编程序逻辑控制器(programmable logical controller)。第一台可编程控制器于 1969 年诞生,并用于美国通用汽车公司的生产线。由于它的动作顺序便于修改,所以改变了过去强电控制系统硬接线逻辑(或称布线逻辑)的旧貌,提高了生产过程的“柔性”。早期的 PC 虽然采用了计算机的设计思想,但是由于当时技术水平所限,所用元器件甚多,磁芯储存器件线路复杂,体积庞大,故应用还不广泛。随着大规模集成电路和微处理机技术的发展和成熟,PC 在近十年来获得了异乎寻常的发展。现代的 PC 几乎无一例外地采用了微处理器作为主控制器,用大规模集成电路作为存储器,使得可靠性、功能、价格、体积都达到了比较成熟和完美的境界。PC 的编程采用人们所熟悉的过程语言,类似原先电气工程师所熟悉的梯形图,用“串”、“并”、计时、计数的概念使计算机语言大众化,因而技术人员非常乐于接受,这也是它得以广泛应用的原因之一。目前世界上几十家著名的电气工厂都在生产 PC 装置,从而成为当代电控装置的

主导,为开发新型机电一体化产品提供了技术保证。

2. 特点

(1) 高可靠性 可编程控制器是一种可靠性极高的工业控制设备。随着器件水平的提高,可编程控制器的可靠性还在继续提高。例如三菱 F1, F2 系列平均无故障时间可达 30 万小时,而 A 系列的可靠性又比之高几个数量级,尤其是近年来开发出的多机冗余系统和表决系统则更一步增加了可靠性。可编程控制从软件和硬件两方面来解决可靠性问题。在硬件方面,首先是选用优质器件,再就是设计合理的系统结构,加固、简化安装,使其易于抗冲击,对印刷电路板的设计加工及焊接工艺都做到严格要求。所有输入输出电路一律采用光电隔离,做到电浮空,无论对于抗干扰还是对于方便接地都大有好处。各 I/O 端子除了采用常规模拟量滤波之外,还加上数字滤波。内部采用电磁屏蔽,防止辐射干扰。采用较先进的电源电路,防止由电源回路串入的干扰。采用合理的电路程序,一旦某模块有故障,在线插拔、调试时不会影响整机的正常运行。在软件方面,设置了警戒时钟 WDT,可编程控制器运行时对 WDT 定时刷新。如果程序出现了死循环,就能立即跳出重新启动,并报警。为避免由于程序出错而错误运行,每次扫描都对程序进行检查和校验,一旦程序出错立即报警并停止运行。对程序及动态数据进行电池后备,停电后运行停止,但有关状态及信息不会因此而丢失。随时对 CPU 等内部电路进行检测,一旦出错,立即报警。采用上述措施后,可编程控制器的可靠性大为加强。

(2) 编程方便,易于使用 可编程控制器采用与实际电路接线图非常接近的梯形图,这种图形编程方式易懂易编。为了进一步简化编程,编程工作集中到了设计思想的本身而不是如何实现设计思想。当今的可编程控制器还针对具体问题设计了诸如步进顺控指令、流程图指令等指令系统。从硬件方面,使用可编程控制器,无论是接线、配置都极其方便,不要自制很多接口电路。

(3) 环境要求低 可编程控制器适用于恶劣的工业环境。

(4) 与其它装置配置连接方便 对于开关量,输入可以是无源触点开关或集电极开路晶体管输出,输出有继电器、可控硅、晶体管等各种不同的形式,可直接接各种不同类型的接触器、电磁阀等。

对于模拟量,只要模拟信号电平在一定的范围内(通常为 $\pm 10\text{V}$ 或 $\pm 20\text{mA}$),就可以按要求自由设置转换特性,而不需要另加电平转换。另外还有运用热电偶直接输入的 A/D 转换器等。此时就连放大器、冷端补偿也是多余的。

对于数据通信,只需同轴电缆和普通 RS-232 或 RS-422 接口即可,不必由用户来考虑波特率及通讯规程等具体的设置问题。

3. 可编程控制器工作原理

可编程控制器是一种工业控制计算机,其核心就是一台计算机。但由于有接口器件及监控软件的包围,因此,其外形不像计算机,操作使用方法、编程语言甚至工作原理都与计算机有所不同。另一方面,作为继电控制的替代物,由于其核心为计算机芯片,因此与继电器控制逻辑的工作原理也有很大区别。可编程控制器的工作过程如图 7.18 所示。

图 7.18 可编程控制器工作过程

(1) 输入处理 程序执行前,可编程控制器的全输入端子的通/断状态读入输入映像寄存器。

在程序执行中,即使输入状态变化,输入映像寄存器的内容也不变。直到下一扫描周期的输入处理阶段才读入这变化。另外,输入触点从通(ON) 断(OFF) [或从断(OFF) 通(ON)]变化到处于确定状态止,输入滤波器还有一响应延迟时间(约 10ms)。

(2) 程序处理 对应用户程序存储器所存的指令,从输入映像寄存器和其它软元件的映像寄存器中将有关软元件的通/断状态读出,从 0 步开始顺序运算,每次结果都写入有关的映像寄存器。因此,各软元件的映像寄存器的内容随着程序的执行并不断变化。

输出继电器的内部触点的动作由输出映像寄存器的内容决定。

(3) 输出处理 全部指令执行完毕,将输出 Y 的映像寄存器的通/断状态向输出锁存寄存器传送,成为可编程控制器的实际输出。

可编程控制器内的外部输出触点对输出软元件的动作有一个响应时间,即要有一个延迟才动作。

4. 可编程控制器的构成

可编程控制器有各种不同的结构,为了简化问题,将以小型可编程控制器为例。

(1) 硬件 可编程控制器的核心是台单板机。它包括中央处理器 CPU,存储器

RAM,ROM,并行接口PIO,串行接口SIO,时钟CTC。它的作用是对整个可编程控制器的工作进行控制,共分两部分:一部分是对系统进行管理,如自诊断、查错、信息传送时钟、计数刷新等;另一部分就是根据用户程序执行输入输出操作,程序解释执行操作等。

单板机中的CPU芯片随机型不同而有所不同。随着超大规模集成电路制造水平的提高,可编程控制器的芯片越来越高档。存储器主要用于存储系统监控程序及系统工作区间,并且用于生成用户环境,其容量的大小取决于系统的工作能力及系统程序的质量。串行接口和并行接口是用于CPU与接口器件交换信息的,它的数量取决于系统规模的大小。定时器/计数器是用于产生系统时钟及用户时钟信息的。

可编程控制器的一个重要特点就是所有输出信号全部都经过了隔离,无论何种形式的输入输出最终都是经过光电耦合口或继电器将信号传入送出CPU。

通常输入有两种形式:一种是直流输入,其输入器件可以是无源触点或传感器的集电极开路晶体管,它又进一步分为源型(共+端)和漏型(共-端);另一种是交流输入,这实际上是将交流信号经整流、限流后再通过光耦传入CPU。

输出通常有三种形式:一种是继电器输出型,CPU接通继电器的线圈,继而吸合触点,而触点与外线路构成回路;另一种是晶体管输出,它是通过光耦合使开关晶体管通断以控制外电路;再一种就是可控硅输出型,这里的可控硅是采用光触发型的。

可编程控制器还有专门的扩展接口,扩展接口是用于扩展I/O单元的,它使可编程控制器的点数规模配置更为灵活。这种扩展接口实际上为总线形式,可以配接开关量的I/O单元,也可配置如模拟量、高速脉冲等单元以及通信适配器等。

可编程控制器本体上通常是不带编程器的。为了能对可编程控制器编程及监控,可编程控制器上专门设置有编程器接口,通过这个接口可以接各种形式的编程装置,还可以利用此接口做一些监控的工作。

为了存储用户程序以及扩展用户程序存储区、数据参数存储区,可编程控制器上还设有存储器扩展口,可以根据使用的需要扩展存储器,其内部也是接到总线上的。

(2) 软件 在可编程控制器中,软件分为两大部分:第一部分为系统监控程序,它是每一个可编程控制器成品必须包括的部分,是由可编程控制器的制造者编制的,用于控制可编程序控制器本身的运行;另一部分为用户程序,它是由可编程控制器的使用者编制的,用于控制被控装置的运行。

系统管理程序是监控程序中最重要的一部分,整个可编程控制器的运行都由它主管。管理程序主要完成控制可编程控制器何时输入、何时输出、何时运算、何时自检、何时通信等,进行时间上的分配管理;同时也完成存储空间的管理和系统自检。

监控程序的第二部分为用户指令解释程序、编辑程序。

监控程序的第三部分是标准程序模块和系统调用,这部分是由许多独立的程序块组成的。各自完成不同的功能,有些完成输入、输出,有些完成特殊运算等。

用户程序是由可编程控制器的使用者编制的针对控制问题的程序。它是用梯形图或某种可编程控制器指令的助记符编制而成的,可以是梯形图、指令表、高级语言、汇编语言等,其助记符形式随可编程控制器型号的不同而略有不同。

5. 可编程控制器的结构特点

可编程控制器不但工作原理与计算机不同,而且为了便于装入工业现场,便于扩展,便于接线,其结构与计算机有很大的不同。通常,可编程控制器的结构分为单元式和模块式。

单元式的特点是非常紧凑。它将所有的电路都装入一个模块内,构成一个整体,这样体积小、成本低、安装方便。由于在一个单体内集中了 CPU 板、输出板、电源板等,对于某一个单体的输入输出就有一定的比例关系。为了达到输入输出点数灵活配置及易于扩展的目的,某一系列的产品通常都由不同点数的基本单元和扩展单元构成。其中的某些单元为全输入或全输出型,单元的品种越丰富,其配置就越灵活。三菱的 F1, F2 系列可编程控制器就属于这种形式。在这两个系列中,有 60 点(36 输入, 24 输出), 40 点(24 输入, 16 输出), 30 点(18 输入, 12 输出), 20 点(12 输入, 8 输出)的主单元和扩展单元;有 12 点全输入和 8 点全输出的扩展单元。

现在在小型可编程控制器中可配置各种特殊功能单元。这些单元有模拟量 I/O 单元、高速计数单元、位置控制单元、凸轮控制单元、数据输入输出单元等,大多数单元都是通过主单元的扩展口与可编程控制器主机相连的。有部分特殊功能单元通过可编程控制器的编程接口相连。还有的通过主机上并接的适配器接入,不影响原系统的扩展。

模块式可编程控制器采用搭积木的方式组成系统。在一块基板上插上 CPU、电源、I/O 模块及特殊功能模块,构成一个总 I/O 点数很多的大规模综合控制系统。这种结构形式的特点是 CPU 为独立的模块,输入、输出也是独立模块,因此配置很灵活,可以根据不同的系统规模选用不同档次的 CPU 及各种 I/O 模块、功能模块。其模块尺寸统一、扩展、安装整齐,对于 I/O 点数很多的系统造型、安装调试、扩展、维修等都非常方便。目前大型系统大多采用这种形式。这种结构形式的可编程控制器除了各种模块外还需要用基板将各模块连成整体。

7.3 智能监视与诊断

7.3.1 智能监视与诊断系统的开发

智能监视与诊断系统的开发属于工程应用性系统开发。它大体要经过三个阶段,即明确设计准则及要求、制定方案和实施方案。为较好地完成系统开发,以下将分别说明在三个阶段中着重应考虑的问题。

1. 明确设计准则及要求

各种系统都有它自己特定的任务和要求,这需要通过现场实地调查研究来确定。但智能监视与诊断系统也将有一些共性要求,这些要求也是设计者需要考虑的或者说应给予重视的准则问题。

(1) 恰当地采用先进技术,使系统取得较完美的操作性能 虽然各种设计任务的

要求高低、层次不同,一般来说应在所要求的档次中,尽量取得较满意的效果。为此不要忽视对一些可以应用(或者是移植)的先进技术手段、工具的采用。由于目前各行业新技术的发展很快,如何综合利用这些新技术,使产品系统有更好的操作性能?采用何种形式的智能化方案?这是设计中应首先考虑的准则。

(2) 标准化、系列化和可扩展性 为了使设备能配套使用以及设备的更新扩展,在设计制定系统方案时,必须了解已有的那些标准化系列化的产品,它们的通用性如何?直接选用标准和系列产品不仅可以给设计带来方便、省时,而且有利于国家现代化工业的发展。同时,设计应考虑可扩展性,使产品留有发展、更新换代的余地。

(3) 可靠性、安全性 任何工业应用的智能监视系统,其可靠性、安全性总是最为突出和最重要的指标。一旦由于不可靠产生异常或事故,将会对整个生产过程产生严重或不可弥补的影响,因而在设计制作系统时,要尽可能采用各种抗干扰、防护措施以及可靠性保障措施。

(4) 实用性、经济效益及高性能价格比 对于应用技术性的系统,其实用性与经济指标是相当重要的。在追求前三项指标的同时,必须考虑到产品系统能否适应当前国内市场的需要,实用性是否大?如何改造就可以使其实用性更强,而又能得到较高的性能价格比,以便能产生更高的经济效益。

2. 制定方案的主要工作

(1) 制定设计任务书 从实际任务到制定出设计任务书之间尚需要经过大量的调查研究。其中主要包含两方面的内容:一是测控对象数字模型对现场各种条件能否精确识别,技术配备,用户的各种要求,国际标准规定等;另一方面是国内外同类系统或类似系统技术指标的比较,应用情况及中外情报资料,专利查询等。在实地调查基础上产生设计任务书,其内容应包括:主题任务、达到目的实际意义、研究内容、技术指标及经济效益等项。

(2) 拟定初步方案 根据设计任务书的要求,研究可能实施的几种设想、规划,并加以比较推敲,选择一种认为可行的,较佳的方案作为初步方案。在初步方案中,对整机系统的结构主要部件环节、技术力量、可能应用开发手段及规划进度等均应有说明。

(3) 方案可行性论证 方案可行性论证是要回答该系统是否可行以及可行的程度。具体地说,要论证贯穿设计工作的指导思想、技术原则是否正确,采用先进技术的必要性及可行性,系统能达到边界条件的最低方案和留有余地较佳方案论证,系统可靠性、性能价格比论证以及系统全局性、整体性诸因素相关的论证等。

(4) 方案制定 在可行性论证得到上一级或用户单位的肯定审核之后,再进一步制定总体实施方案。应明确方案达到的总目标、总体技术指标、总体结构(硬软)方案以及总预算和总进度等。

3. 方案实施过程

方案实施过程中需要涉及系统的总体及各部件的硬、软件设计及分部调试和总体调试等任务。由于以微机作为核心部件的系统,其软硬件常交错在一起,所以硬件尚未齐

备,要在其上运行软件进行调试是不可能的;反之,软件尚未调试好也无法支持硬件的调试。为此系统设计、开发过程中必须讲究一定的方法、步骤以及根据情况配备适当的开发工具,方能顺利进行。现就其中一些主要步骤加以说明:

(1) 确定系统规模大小 系统总体方案确定之后,首先要预估系统软、硬件规模的大小,对硬件核心部件选型、容量、对外的 I/O 口数、通道数、模块等均要进一步核算。

(2) 软、硬件权衡分配 在既定的总体规模中,再进一步权衡,哪些模块用硬件完成,哪些可用软件完成,合理调整好硬、软搭配。从原则上讲,许多硬件功能用软件也可以完成,反之亦然。但是,在不同场合软、硬件将各有特长,要使系统达到较高的性能价格比,必须使应用系统有恰当的软、硬件比例。一般来讲,硬件处理速度快,但应变灵活性小,扩展功能要另添加部件;而软件处理速度慢,但变更灵活性大,添加功能有时只需对软件作适当修正即可。

(3) 硬件部件调试 解决了软、硬分担之后,硬、软件设计基本上可以分开进行,但是相互间的联系是很多的,故也可以交叉进行。先逐个建立硬部件,并进行初试。可使用通用的电子仪器仪表及逻辑分析仪等分别对各部件进行通电前的线路检查、通电试验以及在某种激励信号作用下的响应试验,以验证系统部件功能是否正常。如传感器、模拟通道输入板、模拟通道输出板等都需要经一定的测试校正等试验工作。

(4) 软件模块的调试 软件模块的调试,一般要借助与目标系统相同机种的高中档微机系统,或者采用专门的微机开发系统作为工具。将编写的源程序送入工具微机系统,这类系统必须具有编辑源程序的功能,有语法排错的功能、有交叉汇编或编译功能、模块连接形成目标程序的功能等,以便对软件进行单调初试。

(5) 在线仿真调试 在线仿真调试是系统调试最关键的一步,用以完成实时仿真的软、硬件联调。仿真调试最好使用带有在线仿真器的微机开发系统,从而可以获得高效率的仿真调试,逐步排除故障完成联调。若无微机开发系统也可借助一台功能比较完善,具有一定开发能力,与目标样机系统具有相同的 CPU 机种的微机系统协助开发。

7 3 2 系统硬件体系结构设计

制造过程中,智能监视与诊断系统的体系结构形式,主要由其核心部件——微型计算机硬件及各种输入输出通道的结构所决定。有关输入输出通道的各种结构形式均已讨论,此处着重于应用系统核心部件微机体系结构的讨论。

微型计算机硬件体系结构的选择,主要根据应用系统的规模大小,控制功能性质及复杂程度,实时响应速度及精度等专项指标和通用指标决定。首先根据系统规模及可靠性要求来考虑:对于普通要求的比较小的应用系统可采用单机系统;对于高可靠性系统,即使系统规模不大,但为了可靠,也常采用双机系统;对于规模较小,但系统各部分相对比较集中时,则常采用中、小型微机搭配构成的主从式集散控制系统;对于规模大,且分散在各地的系统应采用多微处理机分散控制系统或分布式数据采集处理系统。

1. 几种常用的单机硬件系统结构

目前由于微机发展及应用的普及,已形成多种档次的通用和专用微机系统,用户可以

用它们通过接口技术和扩展技术直接组成应用系统。这样不仅可以缩短设计周期而且使系统容易达到标准化和通用化。以下介绍几种常用微机系统并比较它们的特色。

(1) 单片机及单板机系统 用单片机或单板机进行适当扩充和接口可构成较小的应用系统。单板机应用早已比较普及,而近期单片机技术得到了迅猛发展。在智能监视系统中单片机技术显示出它的优越性,得到了广泛应用。

由于芯片的集成度高,单片机的优越性主要表现在单一芯片上可以集成微型计算机,性能价格比高,可靠性高,体积、重量均大为减小等方面,因此单片机在数据运算、实现数据通信及多机分布式系统中具有良好的性能。单片机的产品还在飞速发展,其集成度愈来愈高,整机性能愈来愈好,某些产品价格也在降低,是微机使用者值得注目的产品。

(2) 专用工控机(可编程控制器) 80年代以来,日本、美国和德国等国家的一些公司竞相推出各种采用了微型计算机技术的可编程控制器。可编程控制器的结构设计上专用性很强,其产品常由电气控制厂家、公司生产而非计算机厂家。

目前国内可编程控制器的应用也日益广泛,产品已从引进转为自己组装。但生产的厂家尚不多,主要还是靠引进,如日本三菱 F 系列,东芝 EX 系列,美国 GE 公司的 SIRIES 系列,德国西门子公司的 SIMATICS 系列等。各系列产品大体都有小型、中型及大型之分。

小型产品的 I/O 点数小于 128 点,绝大多数用单片机作处理器,其特点是价格便宜、结构紧凑、编程器小巧灵活、结构呈整体密封型、功能仅限于逻辑顺序控制。

中型 PC 产品其输入输出总点数在 128 ~ 1024 之间,这种 PC 的存储容量为 2K ~ 8K,除顺序逻辑控制外大都带有定时、计数运算处理功能,其中高档次产品还具有模拟量控制联机 and 通信功能。中型机代表了 PC 产品的主流。

大型 PC 产品的 I/O 点数可大于 1024 点,存储容量 8K 以上,通常用 16 位或 32 位微处理器,还具有很强的扩充功能和系统组合能力。

目前,各国主要可编程控制器生产厂商还推出了可编程控制器的局部网络,使可编程控制器不仅用于单机,还可以联网。

(3) 通用工控机(工控模板式微机) 可编程控制器由于专用性较强,数据处理能力差,响应慢,并且 PC 产品的标准化程度较差(这是由于各家公司自成体系、相互竞争保密,造成产品不能兼容,软件不能搬用,以至难以形成标准化的结果),所以当前智能监视应用系统结构的另一个发展趋势是模块组合式通用微机。它既适合监视系统,也适合其它应用,故称通用控制机。通用控制机具有系统开放性、兼容性和扩充性好等一系列优点。它更多地应用于性能要求较高、实时性很强、控制任务较为复杂,但工作环境相对比较好的场合。

目前,各种 PC 总线的模板种类不断增加,已达几十种之多。CPU 板主要有 80286, 80386, 80486 和 Pentium 系列,同时还有电源板、I/O 接口板、A/D 和 D/A。此外,还有一些特殊功能模块,如图型模板、网络通信板等。所有这些为系统的构成提供了极大的方便。

2. 多微处理机系统结构

在监视系统任务庞大、监视对象众多或任务复杂的情况下,例如一个复杂的制造过程,用单台中、大型计算机监视并非绝对不可以,然而如将任务、功能分割后,采用多微机系统结构完成,将具有许多优越性。主要表现在:

(1) 增强系统的可靠性 单机控制时,一旦出现软、硬件故障,整个系统将停止工作;而多微机控制时,由于实现功能分担,一机出现故障并不影响全局,而且有的多机系统中,当一台微机出现故障时,系统可立刻进行重构、重排或调用备用机投入,以保证系统正常工作,因而可大大提高系统的可靠性。

(2) 加快处理和速度 多微机的工作既相互独立,又是一个完整的整体。各微机可并行处理是提高处理速度的主要途径。

(3) 性能价格比提高 由于微机的价格,特别是低档微机的价格要比中、大型计算机价格低廉得多,用多台微机构成系统的总价也仍然低廉,而性能又提高,所以性能价格比较高。

(4) 系统结构宜于实现扩充和改造 多处理器系统结构大多采用模块化的标准件。要扩展系统的功能,只要增加相应的模块或微机的台数,就可以重新构成系统;要修改系统功能,也只需换下相应的模块,用新型模块代替即可。

在多微机监视系统中,广泛采用分布式微机系统。分布式微机系统,从总体结构上又可以分为两类;分级分布式结构(集散控制系统)和并行分布式结构。下面简要介绍分级分布式结构。

系统任务分割时可分为几层:上层机负责管理中层机,中间层次机接受上层机的命令,同时又负责管理再下层的微机工作,最低层的微机处理机,则与过程对象直接联系。见图 7.19。

图 7.19 分级分布式系统组成框图

图中纵向联系的计算机有主从结构的特点,而横向联系的微处理机可以有在物理上分散并行处理的特点。管理计算机完成整个生产计划、工艺流程及产品、财务等的管理,

为实现生产过程的静态优化给出数据和指令。监控计算机接收指令,达到过程动态最优化。下层基本处理机单元则分担完成过程现场各种控制任务,并返回信息给上层机。CRT 操作站是显示操作装置,操作人员可以在其上进行操作并通过屏幕显示多种标准画面,观察生产过程、报表、过程监视模拟图等。CRT 操作站还带有人机对话的其它各种输入输出设备、记录、存储设备等。

上述分级分布式系统大多应用在系统各部件虽然分散但还相对比较集中在不远的地域,适合于较大的制造过程。在分级分布式系统中,完成监视任务的计算机在系统中占有极为重要的地位,它的失效会导致整个过程的瘫痪。因而对系统可靠性要求很高的场合,监控级和管理级可采用双机工作,用另一套作为冗余设备,底层处理机采用了真正分散型控制,可分散在现场就地安装,不仅节省电缆,同时减少直接传输中引起的各种干扰,也提高了系统的可靠性。

7 3 3 系统软件设计

系统软件的设计必须根据系统功能的要求,与系统的硬件结构紧相配合,这是监视与控制系统软件设计的特点之一。作为系统软件不论是通用机还是专用机都必须包含两大部件:系统监控软件(或者是操作系统)和系统应用软件。这两种软件从功能上看是不相同的,但从微机软件设计的观点看,还是存在一些共同的要求,如软件标准化、系列化、可塑性(可扩展改造)、可靠性、交互能力好及用户操作使用方便等。

软件标准化是为了系统软件设计时尽量使用已经成熟的标准化、系列化、通用或专用的软件包,以便获得较好的性能,又使系统获得生命力,具有发展前途。但是厂商提供的软件产品,并不能完全满足实际应用的需要,在应用这些标准的软件时,仍需要根据实际系统的结构及要求,配置软件接口并作实际应用的修改,生成新的实用软件。

1. 操作系统监控软件

监控程序属于小系统(或单机系统)低级管理程序(常固化在 EPROM 中),而操作系统是较大的系统或多机的高级系统管理软件,它们都是用来自动化管理内部硬件系统及外设的工具。在实时监视与控制系统中,这种管理还要考虑到实时任务所需的实时设备的协调性,具有多级中断控制能力等,因而称为实时操作系统。实时操作系统主要任务如下:

(1) 微处理机的管理功能 在系统内有多个微处理机或多用户等情况下都存在微处理机管理的问题。其管理方法可为分级调度、作业调度和进程调度等。决定谁有机会占用何级处理机、何时占用、占用多少时间,并且记录微处理机使用的状态等。

(2) 存储器管理 一般计算机系统都有主存、辅存两类。主存也即内存,是计算机能直接寻址访问的存储区;辅存也即外存,是存储的后备空间,容量大。存储器管理主要是合理分配主存。

(3) 外部设备的管理 不论是高级的操作系统或是低级的监控程序都要具备外设管理能力,如管理键盘、面板操作台、显示器或显示终端、打印机及磁盘机等外设。这些外

部设备从操作系统管理的角度看,可把它们区分为独享设备、共享设备与虚拟设备三类。这三类设备在操作系统的统一管理下,按照进程作业进行分配,并完成预定时间的输入及输出操作。在 I/O 完成后,撤消分配,再根据新的进程作业进行再分配。

(4) 文件的管理 文件管理也即负责文件的索取和储存。为了提高文件的存取速度和磁盘利用率,文件按一定的结构排列,如索引式结构、树型结构等。建立好文件系统,用户只要知道文件名字就可以存取相应的文件信息,而无需知道这些文件存放的物理位置。

(5) 系统生成能力 非开发性的操作系统,除上述四种资源管理能力之外,一般也具有一定的系统生成能力。这些能力主要是:可支持的高级语言、汇编语言能力;交叉汇编能力;文件的连接生成能力;动态调试能力;出错处理能力等。

2. 系统应用软件的结构设计

系统的应用软件设计应尽量使用已成熟的功能性的软件包,如数据处理系统软件包、控制系统软件包及某些应用生成软件等,以帮助软件开发。但是作为一个实际应用系统,要能作出满足系统实际要求的软件,其总体结构设计总是必要的。

(1) 应用软件开发过程 应用软件开发过程大体经过以下几个阶段:

分析阶段。分析阶段的主要任务是根据用户要求及实地调查产生出系统描述说明书,确定系统目标、功能、规范等。

设计阶段。在系统任务书的基础上建立软件系统的结构,由总体实施方案划分成若干子系统,确定子系统功能及子系统之间的输入输出联系,进一步确定子系统内的模块,确定各模块功能及各模块之间的输入输出联系,最后确定系统各模块、子系统之间总的逻辑模式图、进程图或结构框图。

编程阶段。编程阶段主要任务是根据模块设计的要求进行程序编制。

调试阶段。进行程序调试、模块调试、子系统调试及总体调试的逐级调试。

运行阶段。在运行阶段中,需将软件系统与硬件系统结合并投入用户环境中实际运行,检验出不适应的部分再进行修改补充。

(2) 结构化程序设计 软件设计,即将模块说明书转换成由某种程序设计语言写成的源程序。所选用的程序语言最好是便于进行结构化程序编程的语言,如 C 语言, PASCAL 语言等。所谓结构化程序设计,就是将各种复杂的流程转换成几种标准形式的基本结构来表达。这几种标准结构是:序列结构、条件分支结构、循环结构和多路开关选择结构,并且每一个基本结构均为一个输入、一个输出。结构化程序设计允许具有递归性,即一个局部框中,又可以由几种基本结构组成。结构化程序设计有许多优点,如程序不易出错、简单明了、容易判错等,故适用于实时监视的程序设计。但是也并非所有的程序都必须用结构化的方法来设计,对某些任务和非结构化的程序语言,使用结构化程序设计反而会带来额外的麻烦时,也就不必勉强使用。

7.4 制造过程的结构模型

7.4.1 输入输出变量的划分

在制造过程中,大部分生产操作的特性由大量动态相互作用的过程变量来表示。这些变量可归纳成两种基本类型:输入变量和输出变量,但还有不同种的输入变量和不同种的输出变量。输入变量有以下三种类型:

(1) 可控输入变量 可控输入变量有时也称控制变量,因为它们在过程中可以改变或控制。在机械加工中,在技术上可以对运行中的速度和进给量进行调整;在化学过程中,可控输入变量可包括流速、温度调整及其它模拟变量。

(2) 非可控输入变量 在运行中可改变,但不能被控制的一些变量称为非可控输入变量。在化学过程中,起始化学原料的变化可以是一个非可控变量,在生产过程中,必须对其作补偿;在机械加工中,如刀具锋利度、工件材料的硬度和工件的几何形状便是非可控输入变量的例子。

(3) 固定变量 加工过程的第三类输入是固定变量。这些是调整的条件,如工具几何形状和工件夹持装置,它们能在各工序间改变但不能在某一工序内作改变。连续化学过程的一些固定输入可以是容器尺寸、蒸馏器的层数以及其它由设备配置所决定的因素。

制造过程中另一个变量的主要类型是输出变量。为方便起见,把输出变量分为两类:

(1) 可测量输出变量 这一类变量的特点是在加工过程中可进行在线测量。在生产过程运行中可测量变量的例子包括流速、温度、振动、电压和功率等。

(2) 性能估价变量 性能估价变量通常与生产过程的经济或与产品制造的质量有密切联系。生产的性能估价变量的例子包括部件成本、生产率、优质产品的产量和质量水平。

7.4.2 结构模型

这些不同的输入、输出变量间的结构关系模型如图 7.20 所示。可测量输出变量由输入变量确定。由性能估价变量表示的生产过程性能由可测量输出变量确定。为了评价生产过程的性能,性能估价变量必须从测量输出变量中计算而得。

图 7.20 制造过程的结构模型

过程控制的问题在于控制一些可测量的输出变量以便获得某些性能估计变量的预期指标,而这要通过对生产过程控制其可控输入变量来完成。

7 5 智能控制策略

7 5 1 反馈控制

在常规的自动控制系统中,最基本的控制系统是简单的反馈控制系统,如图 7 21 所示。在这种系统中,测量元件对被控对象的被控参数(如温度、压力、流量、转速、位移等)进行测量,变送单元将被测参数变换成一定形式的信号,反馈给控制装置,变送单元反馈回来的信号与给定信号进行比较,如有误差,控制装置就产生控制信号驱动执行机构工作,使被控参数的值与给定值保持一致。这种控制,由于被控变量是控制系统的输出,被控变量的变化值又反馈到控制系统的输入端,与作为系统输入量的给定值相减,所以就称为闭环负反馈控制系统,它是自动控制的基本形式。

图 7 21 反馈控制系统框图

在反馈控制系统中,根据实际应用得出的典型控制规律是比例、积分、微分控制规律(简称 PID 控制规律)。当控制系统中仅有比例调节作用时,控制器输出变化为:

$$V_y = K_p e(t) \tag{7.1}$$

根据控制系统理论,要维持系统不发生振荡并有一定稳定的裕度,比例系数 K_p 就不能选择太大;反之,比例系数 K_p 太小,又难免有较大的系统静差,存在静态特性与动态特性的矛盾。

消除静差最有效的办法是采用对偏差信号具有积分作用的控制规律。所谓积分控制规律,即控制器输出的变化量与偏差值随时间的积分成正比。其数学表达式为:

$$V_y = \frac{1}{T_i} e(t) dt \tag{7.2}$$

若有偏差存在,控制器输出会一直由小到大变化,直到输入偏差为零,控制器输出不再变化时,控制器输出速度才为零,即输出不再变化而稳定下来。换句话说,当控制器输出稳定下来不再变化时,输入偏差一定是零。因而积分调节可以消除静差。

单纯积分调节规律控制过于缓慢,因而实际使用中常把比例和积分作用合成,称为 PI 调节规律。PI 调节规律既具有比例调节的快速性,又具有积分作用中那种将偏差积累从而消除静差的优点。PI 调节控制规律输出的变化为:

$$V_y = K_p e(t) + \frac{1}{T_i} e(t) dt \tag{7.3}$$

比例积分系数 K_p 和积分时间常数 T_i 都应恰当选择,否则也会带来不良的控制结果。

控制规律若能和误差变化的趋势有一定关系,即在适当的时机加速或者是减缓控制动作,则可使控制过程得到较稳定的变化,不易产生振荡,而且也不会产生过大的超调现象。这样就需在控制规律中再加入微分环节,即

$$V_y = T_d \frac{de(t)}{dt} \tag{7.4}$$

微分控制规律的作用在于当输入的偏差有变化,控制器就有输出,并且输入偏差变化愈快,控制器削弱其变化的作用就愈大。微分时间常数 T_d 反映了微分作用的强弱。

系统中也不单独使用微分控制规律,常和比例结合,构成 PD 控制规律,或和 PI 结合,构成 PID 控制规律。PD 控制规律为:

$$V_y(t) = K_p e(t) + T_d \frac{de(t)}{dt} \tag{7.5}$$

PID 控制规律为:

$$V_y(t) = K_p e(t) + \frac{1}{T_i} e(t)dt + T_d \frac{de(t)}{dt} \tag{7.6}$$

综上所述,当比例、积分、微分三种控制作用的调节参数选择适当时,有可能充分发挥三种调节规律的优点,得到一种较为理想的静、动态控制效果。

7.5.2 直接数字控制(DDC)

由数字计算机直接对过程进行控制是实现直接数字控制最有效的办法。在 DDC 控制系统中,微机不仅能完全取代模拟式 PID 控制器实现多回路 PID 分时控制,而且不需改变硬件只要改变算法程序,就能有效地实现较为复杂的控制算法。微型机直接数字控制系统又常是分级分布式控制系统底层的现场生产控制机,它是向现代最优化控制发展的阶梯之一。

在直接数字控制系统中,所有的信号处理、显示和控制功能都由一台计算机用数字方式来完成,如图 7.22 所示。由于 DDC 系统的计算机必须执行多种功能,因此,对中央处理器的速度和存储器的要求很高。这些功能包括:多路切换器输入输出扫描、输入信号预处理、数据库的生成、控制算法的执行、工艺及报警画面显示、报表制作、记录和过程优化。中央处理器还必须支持各种后台软件系统,如编译程序、文件系统、文本编辑、数据库建立以及供系统程序员使用的多种实用程序。计算机还必须为它的外部设备(打印机、控制台和外部存储设备等)服务。尽管其它计算机系统也对中央处理器有上述各种要求,但 DDC 系统的要求则是全面和大量的。由于 DDC 系统需要几乎全部组件来维持对过程实行控制,所以任何单个组件的故障都会引起系统失控,其结果导致了系统可靠性和性能的下降。系统的复杂性使得故障的修理工作比常规控制系统困难得多。

图 7.22 直接数字控制系统组成框图

7 5 3 最优控制

上述的各种控制是制造过程控制中属于较为基本的控制方法。当前由于计算机控制及自动控制理论的发展,在自动化程度、控制规律及控制品质上均得到很大发展。许多过去难于实现控制的对象,也有针对性地产生了相应的控制方法。现介绍最优控制方法。

所谓最优化控制,是指在制造过程客观允许的范围内,力求获得制造过程最好的产品质量和最高产量,而能耗又最低的一种控制方法。它的范围可以是一个参数、机组、或者是一个工段、车间和工厂。最优化控制有静态最优化控制和动态最优化控制两种。下面介绍动态最优化控制的若干方法。

1. 线性规划

线性规划研究的问题基本上有两类:一类是在已知原材料及客观限制条件的前提下研究如何完成最多的工作,或使产品达到最好的质量;另一类要完成的任务是预先计划好的,研究如何根据客观所允许的条件,利用最少的能源消耗及最少的原料去完成任务。

线性规划不能直接用来求解非线性问题,但可通过“分段线性规划”的方法得到推广。

动态规划和线性规划虽然名称相似,但两种方法并不相同。动态规划属于动态最优化方法。由贝尔曼所建立的动态规划基本上是多级决策过程最优化的一种方法,故又称为多级决策方法。动态规划在制造工业颇受重视,这是因为大多数制造过程的单元都是由前后相联而成的缘故。动态规划对于这些单元的最优化是将高维最优化问题化为一系列低维最优化问题来处理。

动态规划的基本方法是利用递推关系得出数值解。用动态规划求解一个多单元的实际问题时,由于求数值解需进行的计算次数太多,很繁琐,故要采用解析逼近法,以便在某种程度上减少计算次数。随着新一代计算机的出现,动态规划法会得到更广泛的应用。

2. 多变量搜索法(登山法)

线性规划和动态规划是有限条件下求最大(或最小)目标函数,而且在绝大多数情况下都能得到确定的解答。而登山法则是经过多次反复计算,最后得到的只是一个近似的答案。所谓登山法是一种形象的说法,正如一个人处在山腰上,现在试图爬到顶峰,应如何爬法才能以最快的速度到达山顶呢?很显然,要达到山顶,首先必须确定山顶的方向,否则永远到不了山顶;其次必须选择路线,以最短的路线走到顶峰。

搜索法可分梯度法和模型搜索法。梯度法即根据梯度来确定搜索方向;模型搜索法总的目的是在若干选定的方向上去寻找改进试验的步骤。

3. 调优最优化方法

上述最优化方法认为过程的数字模型是可以获取的,然而工程上许多需要最优化的过程,由于各种原因,很难得到合适的数字模型。图 7.23 所示的压缩机站就是一例,它的目标是要调节原动机的速度,以维持所需的输出压力,且使所需燃料消耗最少。

这是一个约束最优化问题,约束条件就是所要求的输出压力,这类问题用调优最优

图 7 23 压缩机站

化方法可以解决。调优最优化方法的基本思想是用过程本身来求目标函数。在图 7 23 中,最小的量是燃料流量,它可通过仪表直接测量。这样,如果两部压缩机的速度已给出,则所测量的燃料流量就是这种组合的成本函数。然而,在这个例子中,由于约束是要保持输出压力在给定值上,因此,两个速度不能同时独立设定。但是,如果一个速度独立设定,那么为了维持所要求的输出压力,另一个速度就要有相应的值。如图 7 23 所示,这可用一个常规反馈控制器来控制第二个原动机的速度,使输出压力维持在所要求的值上。最优化计算器的功能是确定第一个原动机的速度,以使燃料量为最少,采用的方案是:

在一个适当速度的初始预估值下,测出燃料量的值;

用预定的步长来改变第一个原动机的速度值;

让反馈控制器去调整第二个原动机的速度,使输出压力维持在要求的值上,经过足够的时间,再测量燃料量;

如果燃料量减少,则表示最优化计算器在改变第一个原动机速度的方向上是正确的,并在同一方向再走一步,如果燃料量增加,那么应在相反的方向改变。

7 5 4 自适应控制

本节将讨论自适应控制的一些基本概念及方法。由于自适应控制的研究所涉及的知识领域较广,从事该理论研究的学者很多,加之其本身仍处在不断的发展完善中,所以其定义及分类有很多种,这里选择了两类被广泛认可的类型。

自适应系统能利用可调系统的输入量、输出量或状态变量来度量某个性能指标。根据测得性能指标与给定性能指标的差异,由自适应机构来调整可调系统的参数或综合一个辅助的控制信号,从而使系统的性能指标接近给定指标。这里可调系统应理解为能通过调整其控制信号来调节其性能的、包含被控对象的系统。这个定义适用于模型参考自适应系统。

自适应控制系统必须提供被控对象的当前状态的连续信息,也就是能辨识对象。它必须将当前系统性能与期望的或者最优的性能相比较,从而得出系统趋向最优的决策或控制。这一定义适用于自校正控制系统。以上定义中涉及自适应控制的三大要素:一是

对象信息的在线积累;二是综合有效控制量的可调控制器;三是对性能指标实行闭环控制。

图 7.24 是模型参考自适应控制的框图,它由参考模型、被控对象,常规反馈控制器和自适应控制回路(自适应律)四部分组成。参考模型是在已知系统输入的前提下为达到理想输出而设计的过程模型;自适应律是为使系统的实际输出趋近理想输出所设计的算法,这一算法要求在系统参数未知或变化时由偏差来调节系统的控制器(或直接给出辅助控制量),以尽可能地减小这个偏差 $e(t)$ 。

图 7.24 模型参考自适应控制框图

设计这类自适应控制系统的核心问题是如何综合自适应律。自适应律的设计目前有两类不同的方法:一种称为参数最优化的方法,即利用最优技术搜索到一组控制器参数,使得某个预定的性能指标,例如 $J = \int e^2(t)dt$ 达到最小;另一种设计方法是基于稳定理论的设计方法,其基本思想是保证控制器参数的自适应调整过程是稳定的,然后再使这个过程尽可能地收敛得快一点。由于自适应控制系统一般都是本质非线性的,因此这种自适应律的设计自然采用非线性系统的稳定理论,李雅普诺夫稳定性理论和波波夫超稳定理论都是设计自适应控制系统的有效工具。

第二类自适应控制是以对被控对象模型参数进行在线辨识的自校正控制,如图 7.25 所示。自校正调节系统由被控制对象、辨识器、控制器组成。辨识器对被控对象的参数、状态变量进行在线辨识、估计,并交给控制器,由预先指定的性能指标来综合出系统的最优控制律,以适应不断变化的系统和环境。由于辨识和综合都是在线实时递推进行的,所以可以适应系统的变化。

图 7.25 自校正控制模型框图

这样做是将系统的辨识与控制分开,理论上讲辨识和控制可以分别采用不同的方法。如辨识采用卡尔曼滤波器、最小二乘、最大似然、辅助变量法,控制采用极点配置法、最小

方差法和无振荡控制算法。不同的控制与辨识方法的结合可以组成不同的自适应控制方案。

7.6 制造过程的智能控制

7.6.1 计算机过程监控

计算机过程监控系统的发展与计算机技术的发展密切相关。在过程监控的初创时期,计算机仅仅用在工厂、实验室或其它测试环境中进行数据采集和分析,因而此时的计算机只起到“离线”的应用。在集中式计算机控制系统发展时期,由于在计算机上提供了与过程装置之间的接口,人们开始试验用直接的方法,使计算机与变送器和执行器之间的信号双向传递都不用人工干涉,并获得成功。计算机配备上变送器、执行器和信号连接装置就完全可以实现过程的检测、监视以及对过程的管理和控制了,然而由于计算机与过程装置之间的双向信号流动是通过硬性的物理连接装置,其中流动的信号都是电气信号,因此计算机不可能与现场装置离得太远,所以每台计算机所控制和管理的过程装置很少,多数情况下的应用是单回路控制。在一般的计算机过程监控系统中,主要采用集中式计算机控制系统、直接数字控制(DDC)、集中型计算机控制系统和分层计算机控制系统,下面介绍后两种控制系统的特点和结构。

1. 集中型计算机控制系统

集中型计算机控制系统结构如图 7.26 所示。它把几十个甚至几百个控制回路以及上千个过程变量的显示、操作和控制集中在单一计算机上来实现,即在一台计算机上实现下面的功能:过程监视、数据采集、数据处理、数据存储、报警和登录、过程控制等。此外,还可以实现生产调度和工厂管理的部分功能。

图 7.26 集中型计算机控制系统结构

集中型计算机控制比起常规仪表控制系统有很大的优越性,其控制功能齐全,而且可以实现模拟仪表难以实现甚至不能实现的功能和先进控制、联锁功能等复杂控制。由于单一计算机的高度集中,因此便于信息的分析和综合,容易实现整个系统的最优控制。对于控制回路增删、控制方案的变化,可由软件来改变,提供了应用的灵活性。大量的模拟

仪表盘可用一个操作站 CRT 显示来替代,改善了人机接口。但是,集中型计算机控制应用存在三个主要问题:第一个问题是集中的脆弱性问题。单台计算机控制着几十个甚至几百个回路,把危险也集中了,一旦计算机发生故障,将导致生产过程的全面瘫痪。第二个问题是性能问题。由于生产装置的控制点和运算是很大的,用单台计算机其速度和容量不够,负荷过载;用多台计算机则要解决交换数据和控制信息的困难。第三个问题是开发周期和人力问题。随着计算机的应用面越来越广,控制水平的不断提高,新的要求不断提出,造成开发周期的不断增加。

由于计算机可靠性的提高和引入有效的双重计算机概念极大地提高了整个系统的可使用性。因此,集中型计算机控制系统在小型工厂或生产装置上还有比较广泛的应用。

2 . 分层计算机控制系统

集中型计算机控制系统暴露了集中控制的重大缺陷,促使控制系统向分散化发展。后来出现的过程现场控制与集中显示操作分离开来的分层计算机控制系统,有两种模式:

(1) 计算机、调节器混合控制系统 在原始型分散控制中,各控制回路的每个模拟仪表(单回路)调节器是相互独立的。当某一回路出现故障时,不致影响其它回路的正常工作,系统可靠性高,但难以实现整个系统的最优控制。如果将计算机与模拟仪表结合起来,如图 7 .27 所示的简单结构图,即现场的控制功能由传统的模拟仪表来实现,模拟仪表的信号也送入计算机,计算机显示过程参数,并根据对象的数字模型进行最优化设计,计算最优操作条件,最后以最优工艺参数给模拟仪表设定值。

图 7 .27 计算机、调节器混合控制系统

随着可编程调节器的发展,可编程调节器替代了模拟仪表进而构成监督控制 and 数据采集加可编程调节器的混合控制系统。对于老工厂技术改造,充分利用原有的仪表是可取的技术方案。

(2) 监督计算机控制系统 SCC 在二级系统中如果模拟仪表用一台计算机来代替,实现 DDC,这就是监督计算机控制系统,系统结构如图 7 .28 所示。SCC 由两级计算机组成,第一级计算机与生产过程连接,并承担测量和控制任务,即完成 DDC 控制,因此,要求可靠性高,抗干扰性强,能独立工作。一般选用工业过程计算机作为 DDC 计算机。

第二级计算机是按照生产过程工况,操作条件的变更信息和数字模型进行必要的转换,给第一级计算机提供最佳给定值和最优控制量等各种控制信息。

图 7 28 监督计算机控制系统

我国一些小型过程计算机控制系统常用工业 IPC 为 SCC 计算机,第一级采用 STD 总线工业控制计算机或 PLC 来实现 DDC,两级之间采用串行通信,很适合小型装置的应用。

7 6 2 智能集散型自动控制系统

智能集散型自动控制系统又名分布式计算机控制系统,其实质是利用计算机技术对生产过程进行集中监视、操作、管理和分散控制的一种新型控制技术。它是由计算机技术,信号处理技术,测量控制技术,通信网络技术和人机接口技术相互发展、渗透而产生的,既不同于分散的仪表控制系统,又不同于集中式计算机控制系统。它是吸收了两者的优点,在它们基础上发展起来的一门系统工程技术,具有很强的生命力和显著的优越性。集散型控制系统概括起来由集中管理部分、分散控制监测部分和通信部分组成。集中管理部分又可分为工程师站、操作站和管理计算机。工程师站主要用于组态和维护,操作站则用于监视和操作,管理计算机用于全系统的信息管理和智能化控制。分散控制监测部分按功能可分为控制站、监测站或现场控制站,它用于控制和监测。通信部分连接集散型控制系统的各个分布部分,完成数据、指令及其它信息的传递。集散型控制系统软件是由实时多任务操作系统、专家系统、数据库管理系统、数据通信软件、组态软件和各种应用软件组成。

1 . 集散型控制系统功能分层体系

今天,层次化已成为集散型控制系统的体系特点,使之体现集中操作管理、分散控制的思想。有人把集散型控制系统的层次分成以下四级,见图 7 29。

(1) 现场装置管理层次是直接控制级(过程控制级) 在这一级上,过程控制计算机直接与现场各类装置(如变送器、执行器、记录仪表等)相连,对所连接的装置实施监测、控制,同时它还向上与第二层的计算机相连,接收上层的管理信息,并向上传递装置的特性数据和采集到的实时数据。

(2) 过程管理级 在这一级上的过程管理计算机主要有监控计算机、操作站、工程师站。它综合监视过程各站的所有信息,进行集中显示操作,控制回路组态和修改参数,

图 7 29 集散型控制系统四层结构模式

优化过程处理等。

(3) 产品管理级(生产管理级) 在这一级上的管理计算机根据产品各部件的特点,协调各单元级的参数设定,是产品的总体协调员和控制器。

(4) 工厂总体管理和经营管理级 这一级居于中央计算机上,并与办公室自动化连接起来,担负起全厂的总体协调管理,包括各类经营活动、人事管理等。

2 . 集散型控制系统体系结构中各层的功能

新型的集散型控制系统是开放型的体系结构,可方便地与生产管理的上位计算机相互交换信息,形成计算机一体化生产系统,实现工厂的信息管理一体化。图 7 .30 列出各层所实现的功能。

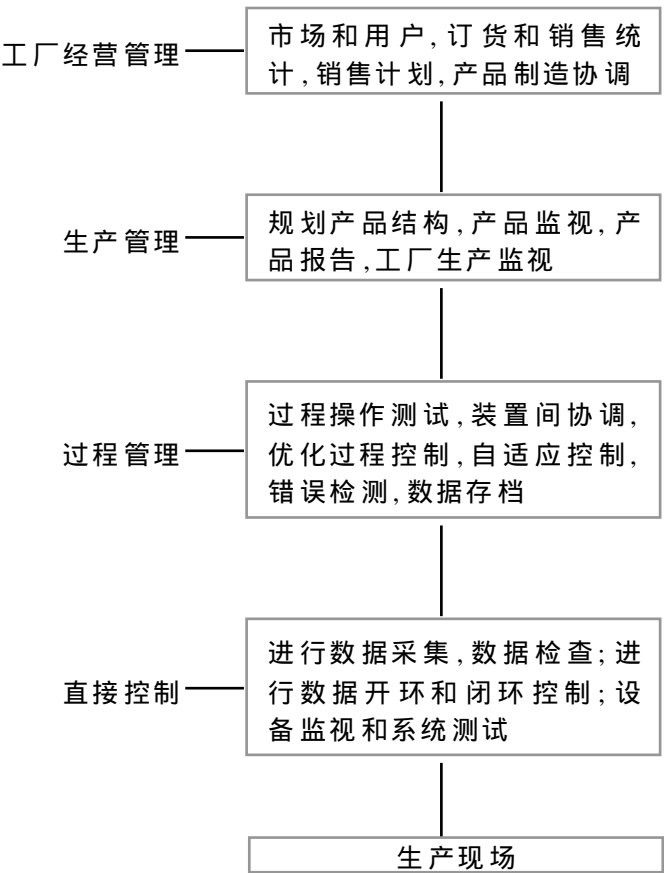


图 7 .30 集散型控制系统结构的各层功能

3. 直接控制级

直接控制层是集散型控制系统的基础,其主要任务有:

(1) 进行过程数据采集 对被控设备中的每个过程量和状态信息进行快速采集,使进行数字控制、开环控制、设备监测、状态报告的过程等获得所需要的输入的信息。

(2) 进行直接数字的过程控制 根据控制组态数据库、控制算法模块来实施实时的过程量(如开关量、模拟量等)的控制。

(3) 进行设备监测和系统的测试和诊断 把过程变量和状态信息取出后,分析是否可以接受以及是否可以允许向高层传输。进一步确定是否对被控装置实施调节,并根据状态信息判断计算机系统硬件和控制板件的性能,在必要时实施报警、错误或诊断报告等措施。

(4) 实施安全性,即冗余化方面的措施 一旦发现计算机系统硬件或控制板有故障,立即实施备用件的切换,保证整个系统安全运行。如友力-2000 集散型控制系统级就是由监测站或(和)控制站组成。可以完成 A/D, D/A 转换、信号调理、开关量的输入/输出,并把采集到的现场数据由 A/D 转换、信号调理或某些信号的直接输入,经过整理、分析,实时地通过高速数据公路传递到上一层计算机中。对于要求控制的量实施实时的调节控制,当发现某一 CPU 板或数据采集板或信号输出板等出现故障就立即向上报告,并根据条件实施切换,以确保系统的正常工作。

4. 过程管理级

过程管理级主要是应付单元内的整体优化,并对其下层产生确切的命令。在这一层可完成的功能有:

优化过程控制,可以根据过程的数学模型以及所给定的控制对象来进行。优化控制只有在优化执行条件确保的条件下方能达到,但即使在不同策略条件下仍能完成对控制过程的优化。

自适应回路控制。在过程参数希望值的基础上,通过数字控制的优化策略。当现场条件发生改变时,经过过程管理级计算机的运算处理就得到新的设定值和调节值,并把调节值传送到直接过程控制层。

采用专家系统,优化单元内各装置,使它们密切配合。这主要是根据单元内的产品、原材料、库存以及能源的使用情况,利用知识库和推理机来协调相互之间的关系。

通过获取直接控制层的实时数据以进行单元内的活动监视、故障检测存档、历史数据的存档、状态报告和备用。

5. 生产管理级

产品规划和控制级完成一系列的功能,要求有比系统和控制工程更宽的操作和逻辑分析功能。根据用户的订货情况、库存情况、能源情况来规划各单元中的产品结构和规模,并且可使产品重新计划,随时更改产品结构,这一点是工厂自动化系统高层所需要的。有了产品重新组织和柔性制造的功能就可以就应付由于用户订货变化所造成的不可预测

的事件。由此,一些较复杂的工厂在三层就实施了协调策略。此外,为了综观全厂生产和产品监视,产品也都在这一层来实现,并与上层交互传递数据。在中小企业的自动化系统中,这一层可能就充当最高一级管理层。

6 . 工厂经营管理级

经营管理级属于工厂自动化系统的最高一层,它管理的范围很广,包括工程技术方面、经济方面、商业事务方面、人事活动方面以及其它方面的功能。把这些功能都集成到软件系统中,通过综合的产品计划,在各种变化条件下,结合多种多样的材料和能量调配,以达到最优地解决这些问题。在这一层中,通过与公司的经理部、市场部、计划部以及人事部等办公室自动化相连,并在企业的经营决策、产品报价和生产计划等需要大量专家知识的领域,采用专家系统,从而逐步实现整个制造系统的智能化。

第 8 章 智 能 装 配

8.1 概述

机器产品一般是由许多零件和部件所组成的。根据规定的技术要求,将若干个零件或部件进行配合和连接,使之成为半成品或成品的工艺过程称为装配。其中把零件装配成部件的过程称为部装,而把零件和部件装配成最终产品的过程称为总装。

装配是整个生产过程的一个重要组成部分,也是整个机械制造工艺过程中的最后一个环节。它在整个机器的生产过程中占有特别重要的地位。对于简单的产品,可以通过若干零件直接装配而成;对于结构比较复杂的产品,为了保证质量和提高装配效率,应根据产品结构的特点,从装配工艺角度出发将其分解为可以单独进行装配的若干单元,即装配单元,如图 8.1 所示。划分装配单元可以更合理地组织装配生产,一般直接进入总装配的装配单元称为部件,进入部件装配的装配单元称为分部件(或组件)。

图 8.1 装配单元

装配工艺过程就是指,将零件或部件按一定的技术要求和顺序安装到规定的位置并与其相关零(部)件联接,最终产生合格产品而采取的一切工艺性操作的生产过程,它包括装配、调整、检测、试验、油漆及包装等工作的全过程。

保证装配精度是保证机器质量的关键之一。其装配精度不仅取决于零件的加工精度,同时也取决于机器产品结构的装配工艺性,装配生产的组织、管理以及装配工艺和技术水平。

目前,装配作业在产品制造过程中所占的比例正在日益增大。据国外统计资料表明,在现代工业化生产过程中,装配作业劳动力占产品制造中的 40% 以上,产品生产的总成本的 50% ~ 70% 是消耗在装配作业中的,而其自动化程度都远远落后于其它作业,成为产品生产过程中的瓶颈。

为了适应市场对产品多样化的要求,柔性自动装配系统在 20 世纪 90 年代得到了迅速发展。同时,人工智能、神经网络在制造业中得到了应用,出现了智能装配专家系统及智能装配机器人等,使装配作业正向着柔性化、自动化、智能化方向发展。这对适应市场变化的快捷响应,提高产品的装配速度和装配精度,保证产品质量,缩短新产品的开发和生产周期起到了十分明显的作用。

8 2 人工装配特点

目前,装配作业仍主要采用以人工装配为主。人工装配是一种领域知识辅助且基于感觉感知的人工作业,主要是依据装配图上的图形、符号、尺寸和文字说明,判认所需的各个零件、零件间的位置关系和配合关系,然后结合装配领域知识和装配人员的经验知识,确定装配顺序、制定装配工艺规程,最终按装配工艺规程将零件装配成机器或部件,完成装配任务。

人工装配是装配操作者用双手在工艺装备辅助下完成的,这不仅需要操作者具有足够的有关装配作业的领域知识和识别装配图纸的能力,如阅读装配图纸、了解装配要求、规划装配路线、选择装配工装夹具、确定中间及总成控制方法等,更重要的是操作者需要具有动手完成装配作业的技艺知识。其技艺知识是一种动态智能,是实时的感觉感知及行动的经验,可以通过示教和实践学到,但难以用文字和语言表述。因此,装配技艺知识是一种由感知、决断、动作、再感知组成的闭环信息过程,只能通过示教学习建立起来的时序感知、决断、动作序列。它的过程模式是:

初始状态(t) 感知 决断 动作

状态($t+1$) 感知 决断 动作

状态($t+2$) 感知 决断 动作

...

目标状态($t+m$) 感知与动作停止。

装配技艺知识的学习和建立大致是个条件反射建立过程。

在人工装配过程中,装配操作者通过多通道感觉系统来感知装配状态。如通过眼睛可感知装配环境和进入装配的零件姿态等;通过施加转矩可感知轴孔紧固状态;通过正反转传动链感知间隙大小;通过推拉滑动副感知运动状态等。通过感知获得有关装配状态的信息,并及时通过人脑的推断、决策,实施正确的装配操作过程。图 8.2 所示为人工装配系统图。

在人工装配过程中,由于每个人的装配技艺知识各异,所以其装配质量也因人而异,例如人的装配技术水平和情绪等因素,往往易导致产品质量变化、性能不一致;而且人容易疲劳,常常处于混乱低效的状态,其必然结果是,人工装配的产品还要求一系列非生产

图 8.2 人工装配系统

性活动(例如检查和返工)来补救其产生的缺陷。同时,装配过程需要判断能力、熟练技巧、灵活性、以及低成本等,而分配在装配工作岗位上的并不能永远具备这种素质。因为人的判断可能出错,人的熟练技巧也是有限的,他们的灵活性不是始终如一的,并且人的体力在工作一段时间后很快就会下降。

可见,人工装配作业的主要特点是:

人能识别表述装配关系的装配图,可根据个人经验和公有知识进行识别、推理、决策和规划等。

人能根据个人技艺知识和借助工艺装备来组织装配作业的实施操作。

人能实时感知装配状态,具有通用、多自由度、灵巧的手脚(执行终端)来实施对装配对象的操作。

装配质量因人而异,取决于操作者的装配知识和个人装配经验以及装配时的装配环境和操作者的情绪因素等。

8.3 智能装配

所谓“智能装配”,就是能自行模拟人自动地实施各种功能的机械装配,其突出特点是在装配过程中具有高度的自适应性,并具有自学习、推理、决策、规划等功能,能按照装配要求实现最佳装配。如图 8.3 所示为智能装配系统的一般构成。

在智能装配系统中,主要有三个重要组成部分:

(1) 知识库 存储装配领域专家的技艺知识和与装配有关的信息。主要有:零件特征信息库、装配技艺知识库、装配规则库等。

零件特征信息库,主要用于存储和输入零件特征信息;装配技艺知识库,主要用于存储专家装配知识和装配经验;装配规则库,存储有效的装配规则、约束条件和策略。

在系统中,装配知识的建立和获取是关键,它主要是从专家的经验归纳出的启发性知识,它需针对新情况不断更新和扩充知识。

(2) 推理、决策机 根据知识库中的装配知识和规则推导最终装配顺序和装配策略,其中包含了推理方式和控制策略。

推理方式有确定性推理和非确定性推理两种,在智能装配系统中,需要处理好装配工

图 8.3 智能装配系统构成

艺过程的非确定知识,因此一般采用非确定推理方式。

控制策略主要有三种:

正向推理,即从已知数据信息出发,正向使用装配约束关系和装配规则求解,推理和导出满足产品性能要求的装配顺序规划。

反向推理,即从最终产品要求出发,按拆卸原则和方法,从拆卸角度去搜索和推理拆卸顺序,然后再通过回溯推理,即可得出装配顺序规划,一般装配工艺的规划和决策采用这种推理方法。

正反向混合推理,交替使用正向推理和反向推理方法,其优点是能结合正、反推理优点,缩短推理过程。

(3) 智能装配机器人 能自行根据装配顺序规划和控制策略去实施自动化装配作业(在 8.3.6 节中详述)。

8.3.1 装配技艺知识表达

由于装配质量对产品性能有直接影响,因此对装配作业一般都有严格的技术要求,如防止表面划碰、控制紧固力矩、保证相对位置精度、保证配合精度、控制间隙大小等。因此装配作业是在一定的装配技艺知识和装配规则约束下进行的工作。

1. 装配规则描述

对一个产品来说,按其联接方式可分为可拆卸产品和不可拆卸产品,而对可拆卸(指无损拆卸)的产品,拆卸是装配的逆过程。在制定装配工艺规划或装配策略时,必须遵循一定的装配原则。如从装的角度考虑,通常在确定装配基准件后,按“先下后上,先内后外,先难后易,先重大后轻小,先精密后一般”的原则;如从拆的角度考虑,可以认为拆是装的逆序,而拆的顺序比装的顺序更容易获得。拆的通常规则是:联接件不拆,被联接件就拆不下来;上下结构的装配件中,上面的零件不拆,下面的零件就拆不下来;左右结构的装配件中,如是由左端装入的,那么左边的零件不拆,右边的零件就拆不下来;同理,如是右端装入的,右边的零件不拆,左边的也拆不下来。可见,联接件与被联接件之间存在约束关系;零件间的相对位置使零件之间存在约束关系。

各种产品的装配都必须按照一定的装配规则和约束条件进行,其规则和约束条件是

多种多样的,下面列举一些拆卸和装配的基本规则。

规则一:可拆卸为零(部)件的产品,一般可以再将零(部)件装配为产品。即可拆卸就可再装配。

规则二:将零(部)件装配成产品后,不一定能将产品再拆卸为零(部)件。即可装配不一定可拆卸。

规则三:可以取产品拆卸工序的逆序列为产品装配工序序列,即先拆下的后装上。

规则四:产品拆卸只要求解除两个相关零(部)件之间的相对有序分布及相对物理关系,一般并没有其它拆卸后的目标要求。拆卸是从高序到低序,而装配则是从低序到高序。

规则五:要在合理范围内,将参与装配的零(部)件装配成部件再参与产品装配,即尽量减少参与装配的零(部)件数目。

规则六:拆卸时导致零(部)件解体的紧固件不宜先拆卸,而宜将由这些零(部)件组成的部件作为整体拆卸。

规则七:装配过程中,前置装配结果要具有抗后续装配动作干扰的能力,即中间装配状态要具有稳定性。

规则八:同一方向的拆卸或装配动作宜集中一次完成。

规则九:任一零部件在实现装配要求的相对有序分布和相对物理关系时,不与任一障碍相干涉,则这一零(部)件就是可装配的。

规则十:产品上任一零(部)件在解除其装配要求时,不与任一障碍相干涉且不会导致产品上其它零(部)件的不稳定,则这一零部件就是可拆卸的。

规则十一:常用组成产品的最大件或与其它组成件物理接触最多的件,如底座、箱体等作为装配基础件,其它多数零(部)件都在此基础上进行装配。

装配规则的知识表达可用产生式规则加以表示,其一般形式为:

IF 条件 1
AND 条件 2
...
AND 条件 n
THEN 结论 1 ,....., 结论 m

对于每条装配规则,如果一定的条件得以满足,那么与它相对应的结论就会产生。目前国内外许多专家系统都是采用这种产生式规则来构造知识库的。其优点是将知识表达成“IF 条件 THEN 结论”这样简单的形式,接近人类思考问题的方式,较容易从专家那里获取领域的知识,且这些规则是独立的,所以很容易被用户检查、修改和扩充,使系统能从一种环境下去适应另一种环境。并可运用置信度的概念,去描述不确定知识的能力等。

2 . 装配状态描述

任何一个问题都可以看成是一个状态空间。问题的求解则是从初始状态在一系列操作控制下,经过若干中间状态,最后达到目标状态的过程。装配状态反映出装配任务执行情况 and 装配过程的状态,包括参与装配的零件状态(如零件空间位置、取向姿态等),装配

机械手状态(如运动轨迹、抓取位置、夹持力、插入力、拧紧力等),配合面接触状态(如接触变形情况、接触质量以及配合的间隙与过盈程度等),装配环境状态(环境变化情况以及所采用的夹具、辅具等)等。

在人工装配过程中,装配状态可通过并行多通道感觉及感知融合获得。如视觉可感知工件的有无、工件的形状、工件之间的相对位置等;触觉可感受接触状态、微幅振动、打滑、表面粗糙度等。人在完成装配作业过程中,除运用视觉、触觉来实时感知装配状态外,还借助于其它感知,如声觉、冷热觉、湿觉等直接或间接感受装配状态(如零部件装配状态不同,在运动时发出的噪声频带分布不同,运动时产生的温度也不同意)。在智能装配系统中,智能装配机器人(在 8.3.6 节中详述)同人一样也具有并行多通道的感觉反馈系统,能自行感知装配状态,并在装配专家知识和经验的支持下,分析、处理感受到的装配状态变量,通过装配状态的反馈来及时调整和控制对装配对象的操作。

装配状态的描述含组成零部件的空间相对有序分布描述和组成零部件的相对约束描述,可通过动态数据库储存。

3. 装配方法

在产品的装配作业过程中,用什么样的装配方法来达到技术要求规定的装配精度,特别是怎样以较低的零件精度达到较高的装配精度,是装配工艺规划的核心问题。因此,采用何种装配方法对装配成本、装配效率和装配精度有较大的影响,它是制定装配工艺规程时需要考虑的重要内容之一,必须根据生产纲领、生产技术条件及产品的性能、结构和技术要求来制定具体的装配方式。在生产上装配方法有完全互换法、部分互换法、选配法、修配法和调整法等。不同的装配方法有不同的特点,但无论采用什么样的装配方法,都必须用装配尺寸链原理来验算产品的装配精度。

完全互换法是指装配时各零件不需经过任何选择、修配或调整,就能达到规定的装配精度。

部分互换法是指装配时各零件不需经过任何选择、修配或调整,使绝大部分的产品达到规定装配精度。

选配法是指装配时有选择地挑选合适零件进行配合,以达到规定的装配精度。

修配法是指装配时选择一组成环作为修配环,并通过修配加工改变修配环的尺寸,达到规定的装配精度。

调整法是指装配时,调整组成环中某一环的尺寸,达到规定的装配精度。

8.3.2 装配任务的描述

对于每一部机器来说,都是由若干个零、部件组成的,这些零件、部件只有通过一定的形式(关系),以一定的顺序联接(装配)起来,才能构成一部完整的机器,并完成规定的功能。因此装配任务的定义和描述包含着两类相关的知识信息,即装配对象间的装配关系,规范而合理的装配任务分解和装配顺序。

对装配对象间装配关系的分析和描述是确定产品装配顺序和制定装配工艺规程的前提。一般在空间相对有序联接的零(部)件之间的主要装配关系有:

零件相对位置关系;
运动副关系;
紧固关系;
联接关系;
传动关系。

相对位置关系一般是指零件间的插入、套入和平面邻接,并且由联接件将零件组装在一起。

运动副关系是指相关零件只存在相对运动并不传递负载。

紧固关系是指相关零部件之间通过紧固力产生的摩擦力维持紧固关系。

零件间的联接关系可以分为可拆联接和不可拆联接,如键、销和螺纹等联接属可拆联接,焊、铆等属不可拆联接。

在智能装配系统中,装配任务的准确描述对建立装配关系模型是十分重要的。产品的装配过程即是解决各种装配关系中的装配约束要求(几何的、工艺的)过程。如零件形状特征的约束(几何约束)、尺寸关联约束(如有配合关系的零件的接合面,其基本尺寸必然相同等)、基准关联约束(如齿轮在装配中必然靠在轴肩上等)等,通过各种约束关系来保证装配关系的正确性。

装配任务的描述可用装配关系树状图来加以表示,如图 8.4 所示。树的层次节点为对应于复杂的装配关系分解成的子装配关系;树的叶节点为可执行的与装配对象有关的装配操作。

图 8.4 装配关系树状图

同一层次的节点之间为与关系,即上层次节点任务的完成取决于本层次的所有子节点任务的完成。

各层子节点的装配关系确定是通过零件间的装配约束知识加以约束和推导的。这正是由于在装配不同零件时存在几何约束和物理约束,某些零件必须在装配其它零件前装配,这种排序知识称为优先约束知识。其约束关系可分为硬约束和软约束,硬约束是装配零件间的几何特性和机械特性决定的约束;软约束是装配经验知识指定的约束。装配顺序的规划和决策也正是通过装配关系和约束关系加以推导和搜索的,这样能提高装配顺序的搜索效率,加快推理速度。从装配关系树状图中也完全反映出了零件的所有装配顺序。

例如图 8.5 所示的产品,它是由零件 1,2,3,4 装配而成的。其装配关系可用树状图

表示为如图 8 .6 所示。

图 8 .5 产品结构

图 8 .6 装配关系树状图

从装配关系树状图中可反映出其装配顺序。图 8 .6(a)为零件 3,4 先装配成部件,再与零件 2 装配,最后与零件 1 装配并将零件 4 插入到底旋入零件 1 中。图 8 .6(b)是将零件 1,2 先装配成部件,再与零件 3 装配,最后将零件 4 插入到底旋入 1 中。可见树状图反映其装配顺序,当加入约束关系即可获得最佳装配顺序,如本例加入稳定性约束和一定装配规则后,很明显(a)优于(b)。

8 3 3 装配顺序规划

在智能装配系统中,最重要的是装配顺序的规划与决策。这种装配顺序不单单是给零件安排一个装配前后顺序的时间规划,而实际问题非常复杂,需要运用领域专家的装配技艺知识、规则和经验加以辅助,通过一定的装配规划演算,才能搜索出一种最佳装配顺序和装配策略。例如图 8 .7 所示的工件,设是由 1,2,3,4 零件按一定的装配关系装配而成的。其装配顺序必须是先将零件 1 与 4 装成(1,4)组件,将零件 2 与 3 装成(2,3)组件,然后将组件(1,4)与组件(2,3)再组装成最后的工件,或者是先将 1,2 零件装成(1,2)组件,3,4 零件装成(3,4)组件,再将组件(1,2),组件(3,4)装成最后工件。

图 8 .7 工件

由此可见,在多种装配策略中,还要根据实际情况和经验知识选择(决策)最佳装配路线。而不同的装配顺序对产品的装配效率和装配难易程度以及装配质量都有一定的影响,因此在智能装配系统中,装配顺序的规划与决策是实现智能装配的最重要的组成

部分。

自 1984 年 Bour jault 提出一个基于问答生成装配优先关系的装配规划方法以来,新的装配规划研究方法不断提出。1987 年 De Fazio 和 Whitney 在 Bour jault 的求解方法基础上,利用问题的归集性把提问问题的数量从 2^N 降到 $2n$ (n 为装配零件间的关联数)。1988 年 M .sphitalni 提出了连接图法(Connectivity Graph)。1989 年 Homen De .Mello 与 Sanderson 提出了割集法,并用 AND/ OR 图(与/ 或图)进行装配工艺路径的描述,以及采用 petri 网图改造 AND/ OR 图等描述方法,以提高装配顺序的搜索速度。各种方法决定了装配任务的描述,关系到装配知识的表示和装配顺序的推理算法。每种方法也都是基于一定的装配知识和规则的。

通常装配顺序的推理和搜索,是根据一定的装配规则和经验知识,并视具体情况得出有利于保证装配精度,便于装配及校正的最佳装配顺序的。可见从构成产品的零件的装配关系着手,按一定的装配和拆卸原则,根据一定的装配知识和方法就可获得装配顺序。下面介绍两种装配顺序建模方法。

1. 基于割集的装配顺序建模方法

装配作业就是把零件按一定的装配关系和约束关系,以一定的装配要求组装在一起。如构成产品 A 的所有零件和联接件的集合为 V ,以 a, b, c, \dots 表示零件。 A 中的所有零件的可能拆卸方向的集合为 E ,则装配关系模型可用有向图 $G(V, E)$ 表示,其中: G 代表有向图, V 表示由零件或联接件组成的顶点集合,对每个要装配的零件 a, b, c, \dots 都是一个顶点。对于联接件(螺栓,销,键等)来说,如共同完成同一功能(联接或固定相同的零件),作为一个顶点,并在顶点信息中说明联接件的个数,如是完成不同功能的联接件,每个都看成是一个顶点。 E 代表相邻零件间装配作用关系的有向弧集合,如 $a \rightarrow b$ 表示零件 a 作用于零件 b ,即在拆卸时,零件 a 优先于零件 b 。

例如从图 8.8 所示的旋阀装配图上看有 6 个零件:排除填料 4、有阀体 1、锥形塞 2、垫圈 3、填料压盖 5 和两个螺栓 6。这里只有两个螺栓是联接件,而且完成同一功能,因此作为一个顶点。所要建立的有向图模型 G 的顶点数为 5。螺栓联接阀体和压盖,插入压盖与阀体螺纹联接,因此从螺栓到阀体和压盖分别有两条有向弧,故旋阀的有向图模型如图 8.9 所示。

Danial F Baldwin 等人将图论的割集算法与 Bour jault 提出的子集规则和超集规则结合,实现了更有效的优先关系约束搜索算法。如在有向图的基础上加入优先关系知识和约束条件,其约束知识表达可采用谓词 $MP(p_1, p_2)$ 和 $NL(p_1, p_2)$, (p_1, p_2) 为装配状态, MP 表示“必须先于”, NL 表示“必须不晚于”。因此 $MP(p_1, p_2)$ 是指零件 p_2 必须在零件 p_1 之前装,有其特殊情况:一是 $MP(p_1, [])$,表示零件 p_1 必须在任何零件装配后装配,二是 $MP([], p_2)$,表示零件 p_2 必须最先装配。这样加入优先关系和约束条件后将得到比较完整反映装配信息的装配优先约束有向图,使有向图具有了丰富的装配顺序规划知识。再采用割集法就可快速推导出产品的装配顺序。

装配割集法概念:如果在装配优先约束有向图 G 中,有一个边集 S 具有这样的性质:
从有向图 G 中删除 S 的所有边,使有向图的秩正好减少 1;

图 8 .8 旋阀装配图

图 8 .9 旋阀有向图模型

S 的任一真子集都不具备上述性质。
则称边集 S 为有向图 G 的割集。由边集 S 将装配优先约束有向图 G 分割成的子有向图
为 G_1 和 G_2 ,即为装配割集。假设被分割成的两个子有向图 G_1 和 G_2 的顶点集分别为 V_1
和 V_2 ,则 V_1 和 V_2 有如下关系:

$$\begin{aligned} V_1 \cup V_2 &= V \\ V_1 \cap V_2 &= \varnothing \\ V_2 &\neq V_1 \end{aligned}$$

由装配优先约束有向图可求得基础装配割集矩阵,对基础装配割集矩阵的所有操作
运算可得到装配割集矩阵。对割集矩阵的每一行作出相应的新有向图,对它进行连通性
计算,可得全部真割集。用装配割集法求解的关键是判断连通性。连通性计算算法如下:

设有 n 个顶点的有向图 G 的邻接矩阵为 R , 则其秩 $\text{rank}(R)$ 与有向图连通性的关系为:

$$\begin{aligned} \text{rank}(R) &= n - 1, G \text{ 图连通} \\ \text{rank}(R) &< n - 1, G \text{ 图不连通} \end{aligned}$$

因此,真割集的判断可由关联性方法计算。设割集矩阵每一行所对应的新有向图为 G^* , 其秩为 $\text{rank}(R^*)$, 则

$$\begin{aligned} \text{rank}(R^*) &= n - 2 \text{ 为真割集} \\ \text{rank}(R^*) &< n - 2 \text{ 为伪割集, 去除} \end{aligned}$$

割集法是根据有向图的拓扑信息,把装配产品有机地分割成子装配割集的,产品装配割集内部的零件是相互连接作用的,利用此特征,采用割集法就能产生装配所有可能的割集,再利用约束条件进行判断,选择合理的子装配割集。在此基础上,再把子装配割集有机地分割成更小的子集,如此分割下去,直至最后每个被分割的零件子集为所需求的单个零部件为止。

应用割集法能自动产生装配产品的全部装配顺序。

2 . 基于知识 **Petri** 网的装配规划建模方法

(1) 与/ 或图 (AND/ OR 图) 表示 S . Luiz, Homem De . Mello 提出了用 AND/ OR图(与/ 或树图)表达装配关系的方法。例如,假设某产品由盖 1、筒 2 和柄 3 组成,它有两种装配顺序:先 1 与 2 装配,再装 3;先 2 与 3 装配,再装 1。故产品装配与/ 或图如图 8 .10 所示。图中用两有向弧及其间的小圆弧表示一次子装配,两个有向弧及所指的两个零(部)件之间为“与”关系。表示零、部件或产品的方框,称为结点,编号为 1 ~ 6。

图 8 .10 产品装配与/ 或图

图 8 .11 与/ 或树图

由 6 个结点可构造 6×6 的邻接矩阵,将装配关系加入到图论中可直接构造如图8 .11 所示与/ 或树图。圆代表树节点,圆内数字表示零件号或组成部件(产品)的零件号。圆外

数字表示与邻接矩阵中相同的结点号。

如将上述已知的与/或图用知识 Petri 网图加以重构,并采用迭代算法,可获得最佳装配规划。此建模方法具有智能性和适应性强,规划决策速度快和建模自动化程度高等特点。

(2) 知识 Petri 网图 知识 Petri 网结构是一个七元组:

$$P_N = (P, T, I, O, I_c, A_s, \mu)$$

其中(P, T, I, O, μ)与基本 Petri 网相同,其标识 μ 能随零件流而注入。 A_s 和 I_c 与知识有关,其中 I_c 是网上的控制流,由抑制弧和允许弧构成。图 8 .12 所示为知识 Petri 网图,它是在基本 Petri 网图的基础上增加控制流。图中 p_s 为抑制弧位置,用抑制弧引向转移 t_1 和 t_2 ,起流向控制作用。运行开始时 $\mu(p_s) = 1$,所以当转移 t_1 发射后, $\mu(p_s) = 0$,禁止了转移 t_2 的发射,反之亦然。所以 p_s 代表一定的知识,控制 t_1 或 t_2 的发射。

知识 Petri 网的运行规则与基本 Petri 网的运行规则基本相同。

图 8 .12 知识 Petri 网图

将图 8 .11 所建立的与/或树图采用知识 Petri 网图重新构造,得到图 8 .13 所示的装配顺序规划网图。位置(圆)内的数字表示结点的零件或组件号(注意,这儿不表示令牌数)。位置 p_i 的下标 i ,为结点号。但为了便于说明,在网图中用“ ”来表示结点相同的另一结点。

装配规划网 P_N 的运行过程,即是规划决策过程,可采用事件并发异步调度法和转移连续扫描法来运行。零件到达装配站,一次子装配完成等都表示一个事件,都将异步触发 P_N 。当第一个事件发生时开始计时,运行前,位置中的令牌数除 p_s 中的为 1 外,其余均为 0。设零件 1 首先到达装配站,计时开始,位置 p_1 和 p_4 中的令牌数均变为 1,扫描 $t_1 \sim t_4$,均不满足发射条件;一段时间后,零件 2 到达, p_5 和 p_5 中的令牌数均变为 1,此时转移 t_3 满足发射条件,零件 1 与 2 开始装配。装配完成时,使 p_2 的令牌数变为 1, ...。当运行到 p_1 的令牌数变为 1 时,运行结束,获得最佳装配规划。

知识 Petri 网具有能恰当处理并发现象(因果上的不依赖性)和非确定性(选择)现象的能力以及用网状图表示系统模型的方法。将装配过程视为一个离散事件系统是很自然的,采用 Petri 网建模装配过程以及离散事件控制装配过程是很有意义的。

图 8.13 装配顺序规划网图

8.3.4 基于特征的装配方法

1. 装配特征

基于特征的装配是一种产品的装配建模方法,在本书第 4 章中,我们对产品的特征建模已进行了一些介绍,现在我们就零件与装配有关的特征再加以简单介绍。在智能装配系统中,使用人工智能技术,利用专家丰富的经验和装配知识来分析产品装配过程所需要的特征信息,从而实现产品的智能化装配。

零件特征是描述参与装配的零件尺寸和几何形状信息的。在机械装配中最关心的是零件表面,所以可采用关键面特征来表示零件特征。所谓关键面就是能表征零件形状特征的表面,如圆柱体的关键面是圆柱面,圆锥体的关键面是圆锥面,以及圆球面、螺纹面、平面等。关键面的表示要按顺序进行,即从零件的一个端面到另一个端面,且外表面的表示次序优先于内表面。建立这样的零件特征知识库,可保存所要装配的机械零件尺寸和几何形状信息,且建库方便可行。

装配特征是用于表达装配过程中需用的信息的。主要有:

(1) 联接特征 产品的联接方式有可拆卸联接和不可拆卸联接。可拆卸联接的特点是相互联接的零件拆卸时不损坏任何零件,且拆卸后还能重新装在一起,如常用的螺纹联接、键联接和销联接等为可拆卸联接;不可拆卸联接的特点是被联接的零件在使用过程中是不可拆卸的,如果拆卸就会损坏某些零件,常见的有焊接、铆接和过盈联接等。一般拆卸都是从联接件入手,通常拆卸顺序比装配顺序更易获得,装配是拆卸的逆过程。

(2) 装配精度特征 装配精度包括距离精度、相对位置精度、相对运动精度以及配合表面的配合质量和接触质量,将直接影响最终产品的质量和性能。

距离精度是指为保证一定的间隙、配合性质、尺寸要求等相关零部件间的距离尺寸精度。相对位置精度是指相关零件的平行度、垂直度和同轴度等方面的要求。相对运动精

度是指产品中有相对运动的零部件间在运动方向上的精度。

(3) 配合特征 是指两个相关零件配合表面之间达到规定的配合间隙或过盈的程度,它包括间隙配合、过渡配合和过盈配合。它反映配合性质和配合质量,影响装配精度。

(4) 基准特征 装配过程中的定位基准和装配基准。通常装配基准有孔、轴、槽、台阶面等。

(5) 约束特征 描述零部件间的装配约束关系以及几何干涉等。

(6) 装配组织形式 装配的组织形式对装配效率的高低和装配周期的长短、装配顺序的安排均有较大的影响。其主要形式是固定式装配和移动式装配。需根据产品的结构特点、生产类型等确定采取何种组织形式。

(7) 辅助信息 包括装配专用工夹具、装配车间、环境等信息。

2 . 基于特征的装配方法

基于特征的智能装配专家系统如图 8 .14 所示。

图 8 .14 基于特征的智能装配专家系统

系统主要由装配知识(特征)库、专家系统、装配顺序规划模块等部分组成。根据产品的功能和装配要求,进行特征抽取,并分析各特征及其之间的关系,划分子装配单元,然后根据特征进行匹配搜索,规划装配顺序,确定装配动作类型,实施装配操作。

8 3 5 柔性自动装配系统

1 . 柔性自动装配系统概述

计算机集成制造系统(CIMS)是现代制造技术和计算机技术发展的产物,它是利用计算机和信息技术把全部生产过程从经营决策、生产管理到设计、工艺、计划、制造、装配、检验、发货等过程综合和协调为一体的闭环系统。它能充分、有效地利用企业的全部资源,对市场需求作出快捷反应,使生产经营过程实现优化、柔性化和适时化。CIMS 是 21 世纪工厂的生产模式,而柔性自动装配系统(flexible automated assembly system, FAAS)是 CIMS 中的一个重要组成部分,它对提高 CIMS 的生产率,发挥 CIMS 的效益有很大的影

响。所谓柔性自动装配系统是一种计算机控制的自动装配系统。它把原材料和购入的零件转变成具有所需功能的产品。

当今,机械制造业竞争激烈,且产品呈多品种、小批量的特点。用户对产品的需求在不断发生变化,对产品的质量和交货速度要求越来越高,柔性自动装配系统正是为了适应这种市场对产品需求的动态要求而产生的。

2 . 柔性自动装配系统的工艺规划

柔性自动装配系统的功能就是要把零件和材料处理成产品,其目的在于提高产品的装配生产率,获得高的装配速度和装配精度,降低装配成本,稳定与改善产品质量,减轻劳动强度以及取代特殊条件下的人工装配劳动。由于柔性自动装配系统具有柔性,所以能在装配系统允许的功能、功率和几何形状许可的范围内处理任何产品。它的优点还在于可以根据需要和变化而作出迅捷反应,如根据需要加入或去掉一些具体装配环节。例如图 8 .15 所示的产品 A,它由 p_1, p_2, p_3, p_4 四个零件组成。对传统的装配工艺规划来说,它需要在三个串联的装配工作站上进行,如图 8 .16 所示。

图 8 .15 产品

图 8 .16 传统装配系统

在柔性自动装配系统中,装配工作站的灵活性可以使多种装配操作在一个装配工作站上进行,因此在典型的柔性自动装配系统中,串联的装配工作站的数量远远小于传统的装配系统。如图 8 .17 所示,可由两台平行的柔性装配站组成,产品可在两台装配工作站的任意一台上装配。

为了使产品的装配效率高,各种零件和部件必须能在正确的位置,正确的时刻,按正确的空间姿态和正确的取向送到装配工位上。柔性自动装配系统的特点是有一条灵活的物料搬运系统,装配件能自由地选择路线,由一个装配站送至另一个装配站,图 8 .18 所示为一种柔性自动装配系统。

当待装配的零件由柔性制造系统(FMS)加工后,依次由输入传送带传送到装配站装配。传送带是断续传送零件的,当某个零件被传送带传送到装配站时,传送带暂停,此时

图 8 .17 柔性装配站

图 8 .18 柔性自动装配系统

装配工作站的装配机器人对零件进行识别,并取走待装配零件。装配作业是在两个具有视觉识别系统的装配机器人的配合下完成的。

装配工艺规划决定装配机器人的操作动作和步骤:一是将取下的零件与已在装配站的零(部)件立即进行装配;二是将取下的零件暂放在缓冲区内,待下一个零件送到后再装配或从缓冲区中选取零件装配。

3 . 柔性自动装配系统的产品设计要求

人们已认识到要提高装配生产率、降低装配成本、实现自动化和智能化柔性装配,必须在产品设计阶段就充分考虑产品的装配要求,使其便于装配。为此便于装配的设计准则(design for assembly)被提了出来。该准则包含两个重要内容:一是尽可能地减少产品中单个零件的数量;二是改善产品零件的装配工艺性。因此在柔性自动装配系统中装配

的产品,除应满足每个零件便于制造外,还必须考虑其装配工艺性。所谓装配工艺性就是装配时易于保证装配精度和装配生产率,使装配简单、可行,适于装配机器人的自动操作。

为此设计适于柔性自动装配系统的产品应该遵守以下一些基本设计原则:

使产品的零件数最少。在装配精度相同的情况下,组成产品的零件数愈少,对零件的加工精度要求愈低,工艺性就愈好。

在保证一定的装配精度条件下,零件的互换性好,装配工作可大为简化。因此在设计时应尽可能按完全互换的原则规定有关零件的尺寸和公差。

要有正确的装配基准。应力求使装配基准能够用零件加工时的定位基准,基准重合能达到更高的装配精度。

尽可能将产品设计成能一层一层装配,这样便于装配和拆卸。

零件的结构应具有一定的柔顺性和对称性,使其易于定位和导向(如采用倒角、锥度等),使装配简单易行。图 8 .19 所示为结构柔顺性。

在装配时尽可能不要旋转或抬起部件(或组件)。抬起或旋转部件会使夹具和装配机器人的手爪复杂,机器人的自由度多,装配周期加长。

紧固件尽可能同一,避免采用大小和类型不同的紧固件。

图 8 .19 结构柔顺性

采用柔性自动装配线带来的效益是巨大的,美国 Xerox 公司为生产小型复印机设计的高速柔性自动装配线建成后已取得如下效益:

减少原料库存量 60% ;

由于能及时地将零、部件送到装配站,所以制品库存量减少了 40% ;

物料搬运的劳动力减少 40% ;

今后更新产品时不需新的搬运系统,因而设备投资可节约 80% 。

柔性自动装配技术已成为机械制造业中的一项重要技术,由于它的巨大应用潜力而受到愈来愈多厂家的重视。据统计在美国销售的装配系统中,柔性装配系统所占的比例也在逐年增加,已从 1980 年的 15% 增长到 1990 年的 45% 。

在柔性自动装配系统中,装配机器人起着至关重要的作用,它已成为柔性自动装配作业系统的主体和核心。

8 3 6 智能装配机器人

为了适应产品多样化和小批量的特点,近年来,对具有柔性的自动化装配系统的需求

日益增多。为满足这种多品种的快捷生产需要,增强企业的市场竞争能力,装配机器人技术得到了迅速发展。

在国外,一些企业的装配作业已大量采用装配机器人来操作。根据日本产业机器人协会的统计,到 1993 年年底,全世界已投入运行的工业机器人有 61 万余台。日本机器人拥有全世界总数的 60%,而日本机器人总拥有量的 96% 用于制造业,其中用于装配作业的机器人占总数的 30%,主要用于汽车等生产行业。

为扩大装配机器人的适用范围,装配机器人正向着智能化方向发展。所谓智能装配机器人就是指人给出某种具体的装配作业指令后,装配机器人能认识装配工作环境、装配对象及其状态,能根据自行理解和判断决策出装配方法,制定出装配顺序规划,实施装配作业。在装配过程中还具有随装配对象变化而适应其工作环境的功能。智能装配机器人的突出特点是具有识别、定位、检测、决策、规划、补偿等多种功能,能适应高速和高精度装配,这对提高产品质量、缩短新产品的开发和生产周期起到了十分明显的作用。

智能装配机器人同人一样应具有视觉、触觉、听觉、嗅觉等多种感觉系统,通过各种感觉感知装配对象和装配环境,然后决策出最佳装配顺序规划,去实施装配操作,使所需的被装配零(部)件按正确的顺序装在正确的位置上,并获得装配成功。

图 8.20 所示为智能装配机器人的系统组成。它主要具有感觉反馈系统、信息处理系统以及决策规划能力,能实时检测和具有故障自动诊断能力。

图 8.20 智能装配机器人的系统组成

在智能装配机器人的感觉系统中,视觉是最有力的获取装配信息的感觉,它使装配机器人能识别装配对象和装配环境(如能识别零件的形状和随意放置的零件的空间姿态和位置等),确保零件的正确空间位置和取向。视觉也是综合性最强的感觉,它在单位时间内获得的信息最多,它能实时反映装配作业过程,也能检测装配作业结果,能检查任务是否完成、有无错误,并能检查出潜在的问题。因此视觉是实现装配机器人智能化的最重要的感觉系统之一。触觉也是智能装配机器人用以感知装配过程和装配状态的感觉系统,可分为接近觉、接触觉、压觉、力觉和滑觉。装有触觉系统的装配机器人能通过触觉感知对物体接近、接触、跟踪、握持、移送、插入或组装等动作状态。也可感知出装配预紧力大小,装配松紧程度等,使装配过程得以正确实施。

图 8.21 所示为日立制作所中央研究所研制的智能装配机器人 HIVIP-MKI,它具有

视觉识别和检测信息系统,可以通过视觉系统中的一只眼睛观看装配图(包括识别零件形状、个数、零件之间的组装方式等),用另一只眼睛寻找工作台上摆放的一堆零件,辨认它们的位置、姿态和尺寸等。根据这些信息,通过运用装配技艺知识、规则推理和决策装配顺序、运送方式,并控制其手臂、手爪等做相应操作,将零件装配成图纸要求的产品。

图 8 .21 智能装配机器人 HIVIP-MKI

在该智能装配机器人 HIVIP-MKI 中,赋予了模式识别和决策的功能,从而只要给它发出接近于人的表现形式(如装配图)的指令,它就可以自行决策达到目标的具体顺序和步骤。

社会需求的日益增长和技术进步,特别是多品种、小批量生产方式以及为提高产品质量及生产效率的生产工艺需求,将都是推动智能装配机器人发展的直接动力。可以预测:随着柔性自动装配系统的进一步发展,企业对智能装配机器人的需求将一定会有较大幅度的增长。

8 3 .7 虚拟装配 VA(virtual assembly)

在激烈的市场竞争下,产品的开发、制造既要质量好,又需要快捷。在这种态势下,并行工程概念被引到产品生产中来。并行工程是一种系统的集成方法,它用并行方法对产品及其相关过程(制造过程和支持过程)进行设计,即产品开发人员从产品设计阶段就需要考虑其制造和装配。并行工程的思想就是要产品设计必须是面向制造的设计和面向装配的设计,使设计出的产品不仅能易于加工,而且能够容易地、经济地进行装配,用以提高装配效率,减少装配时间和装配成本。

对 21 世纪现代化制造厂来说,其明显的特征就是高度柔性化、集成化和智能化,追求的目标都是提高产品的质量及生产效率,缩短设计、制造及装配周期,降低生产成本,最大限度地提高制造业对市场的应变能力,满足用户需求。

面向装配的设计,一方面是以装配集成信息模型为基础,将 CAD/ CAPP/ CAAPP (computer aided assembly process planning)集成起来,使其不但具有加工工艺规程的设计能力,而且具有进行装配操作工艺规程的设计能力,更要有适合于装配机器人自动装配

的装配方式的工艺规程设计能力。另一方面是要在产品设计阶段能实现零部件的虚拟装配。所谓虚拟装配——装配过程的虚拟实现,就是在人工智能和装配工艺知识库的支持下,对所设计的零件进行全过程的装配模拟仿真,从而对该装配件的可装配性作出评价。

虚拟装配技术是由多学科知识形成的综合装配系统技术。其本质是以计算机支持的仿真技术为前提,在产品的设计阶段,实时地、并行地对产品的装配生产全过程进行模拟,预测产品性能、产品的可制造性以及产品的可装配性,从而更有效地、更经济地、柔性灵活地组织生产,使工厂和车间的设计与布局更合理、更有效,以达到产品的开发周期和成本的最小化,产品设计质量的最优化,生产效率的最高化。

虚拟装配是装配作业在虚拟环境下的映射,是装配作业的模型化、形式化和计算机化的抽象描述和表示,它不消耗现实资源和能量,所装配的产品是可视的虚拟产品,具有真实产品所必须具有的特征。

虚拟装配能实现零部件在计算机屏幕上的点到点的装配仿真操作,即零件直接到位的装配操作,并进行干涉检查,如发现被装配零部件与其它零部件或装配环境有干涉现象,或者部件在送往装配位置的过程中有干涉现象,则返回重新修改零件的设计。同时能实现装配工艺的智能设计和决策,能对智能装配机器人进行模拟仿真,使设计者能够在设计阶段了解其所设计的产品的装配工艺过程,通过装配机器人的模拟装配过程显示,了解该产品的可装配性,检验装配过程的可靠性以及装配误差,并以此为依据,改进产品的设计,使设计在满足性能要求的前提下,减少装配成本。

实现产品的虚拟装配是虚拟实现技术在并行工程中应用的一个重要组成部分,是目前国际上刚刚触及到的前沿领域,也是制造业实现并行产品设计与制造的生产模式的关键技术之一。

虚拟装配概念是全新的概念,在这新的技术领域,有许多技术关键尚需进一步研究和开发。这些技术关键是:

- 确定零件之间的安装和配合的概念及其表达方法,以及计算机模型的处理方法;
- 确定零件之间的装配约束关系及装配拓扑关系的描述、建立和处理方法;
- 确定零、部件之间装配、拆卸的计算机操作方法;
- 确定装配特征集的数据结构及表达方法;
- 确定装配集成信息模型的建立、显示及操作方法等。

参 考 文 献

- [1] ohlenius, G . and Kjellberg, T . Artificial Intellegence and Its Potential Use in the Manufacturing System .Annals of the CIRP, 1986,35 (2) : 425—432
- [2] 雷鸣,杨叔子 . 美日重视发展智能制造技术 .机械与电子, 1994 (6): 12—14
- [3] 李伯虎,王纬,陶克强等 . 计算机集成制造系统(CIMS)约定、标准与实施指南 .北京:兵器工业出版社, 1994
- [4] 吴季良 . 柔性制造系统实例 . 北京:机械工业出版社, 1989
- [5] Monostori, L . A Step Towards Intelligent Manufacturing . Annals of the CIRP , 1986, 35 (2): 425—432
- [6] Edmund, C . Payne , Robot, C .Mcarthur . Developing Expert Systems .New York:John Wiley & Soons , 1990
- [7] Robert, I .Levine . AI and Expert Systems . New York: McGraw-Hill,1993
- [8] Wiss,S .M .& Kulikowski, C . A . A Practice Guide to Designing Expert Systems . New Jersey :Row-man & Allanheld ,1984
- [9] Waterman,D .A . A Guide to Expert Systems . New Jersey: Addison-Wesley, 1986
- [10] artric Henry Winston . Artifical Intelligence . New York: Addison-Wesley Company, 1977
- [11] 冯博琴 . 实用专家系统 . 北京:电子工业出版社 , 1992
- [12] 叶景楼 . 人工智能专家系统程序设计 . 沈阳:辽宁大学出版社, 1992
- [13] 吴鹤龄, 马建峰, 李仕文 . 图形专家系统工具 GEST . 北京:北京理工大学出版社,1993
- [14] 范牧昌 . 人工智能在组合夹具设计中的应用 . 机电工程,1994 (4) :5—6
- [15] 钟义信,潘新安,杨义先 . 智能理论与技术 . 北京:人民邮电出版社, 1992
- [16] 陆汝钤 . 专家系统开发环境 . 北京:科学出版社,1992
- [17] 黄可鸣 . 专家系统导论 . 南京:东南大学出版社,1988
- [18] 林尧瑞 等 . 专家系统原理与实践 . 北京:清华大学出版社,1989
- [19] 李桂青,罗持久 . 工程设计专家系统的原理与程序设计方法 . 北京:气象出版社,1991
- [20] 贾焰,王志英 . 知识系统原理与技术 . 长沙:国防科技大学出版社,1993
- [21] 杨行峻,郑君里 . 人工神经网络 . 北京:高等教育出版社,1992
- [22] 庄镇泉,王煦法,王东生 . 神经网络与神经计算机 . 北京:科学出版社,1992
- [23] 曹焕光 . 人工神经元网络原理 . 北京:气象出版社,1992
- [24] 冯天瑾 . 神经网络技术 . 青岛:青岛海洋大学出版社,1994
- [25] 宿延吉 等 . 神经网络理论及应用 . 哈尔滨:东北林业大学出版社,1994
- [26] 安得鲁·库曼克著 . 杨静宇,陆际联译 . 智能制造系统 . 北京:清华大学出版社,1993
- [27] 刘有才,刘增良 . 模糊专家系统原理与设计 . 北京:北京航空航天大学出版社,1995
- [28] 焦李成 . 神经网络系统理论 . 西安:西安电子科技大学出版社, 1990
- [29] 焦李成 . 神经网络的应用与实现 . 西安:西安电子科技大学出版社,1995
- [30] 罗发龙,李衍达 . 神经网络信号处理 . 北京:电子工业出版社,1993
- [31] 焦李成 . 神经网络计算 . 西安:西安电子科技大学出版社,1995
- [32] 何明一 . 神经计算原理 . 语言·设计·应用 . 西安:西安电子科技大学出版社,1992
- [33] 黄国建,王建华 . 提高人工神经网络 BP 算法收敛速度的几种方法 . 华东船舶工业学院学报 .1994,

- [34] Hopfield, J.J. . Neural Networks and Physical Systems with Emergent Collective Computational Abilities . PNAS ,1982 ,79 :2554—2558
- [35] R .Hecht-Nielsen .Neurocomputer Applications .Neural Computer,1988 , 445—453
- [36] Kohonen, L .Introduction to Neural Computing .Neural Network,1988,1(1):3—16
- [37] Pao, Y .H . A Daptive Pattern Recognition and Neural Networks .NewYork: Addison-Wesley,1989
- [38] Henseler, J . Connections, Neurons and Activation .The Netherlands: Met lit .opg ,1993
- [39] Wang jianhua etc .New Function of Cellular Neural Network: Image Reconstruction and Symmetrization .In: Proceedings of IEEE Inter .Conf .on Computers, Communication and Automation .Beijing: Inter .Academic Publishers, 1993,934—937
- [40] Wang jianhua etc .Fuzzy Knowledge-based Artifical Neural Network and Its Application .In: Proceeding of IEEE Inter .Conf .on Neural Network and Signal Processing, Nanjing: 1995, 593 ~ 595
- [41] 曾芬芳主编 . 计算机辅助产品造型技术基础 . 北京: 学苑出版社, 1994
- [42] 白英彩, 曾芬芳等 . CIMS 教程——生产综合自动化系统 . 北京: 学苑出版社, 1993
- [43] 曾芬芳 .船舶照明 CAD 系统的特征建模 .计算机辅助设计与计算机图形学学报 .1995, (3)
- [44] 曾芬芳 .基于参数化的船舶电气自动绘图系统设计 .造船技术 .1995, (3)
- [45] 曾芬芳 .基于参数化设计的轴承 CAD 系统 .华东船舶工业学院学报 .1995, (7)
- [46] Montreuil, B . Requirements for Representation of Domain Knowledge in Intelligent Environments for Layout Design .Computer-Aided Design .1990,22(2):97—107
- [47] Francois Guéna, Knhaldoun Zreik .Analogy , Exploration and Generalization: Three Activities for Knowledge-Based Architectural Design System .CAAD future 's93, Flemming U and Van Wyk S (Editors) ,Elsevier Science Publishers B .V ., 1993, 255—272
- [48] Bloom, G .L .Justifying the Use of CAD in Architecture: Taking Advantage of Representation Requirements .Computer Graphics World .March 1982, 79—80
- [49] 庄越挺, 潘云鹤 .智能 CAD 方法评述 .计算机研究与发展 .1994,31(4):1—7
- [50] Galle, P .Abstraction as a Tool of Automated Floor-Plan Design .Environment and Planning B: Planning and Design ,1986,13:21—46
- [51] Hifi, M .Study of some Combinatorial Optimization Problems: Cutting Stock, Rectangular Packing and Set Covering Problems .Ph .D .Thesis, University of Paris 1 Pantheon-Sorbonne, French .1994
- [52] Prasad, Y .A Set of Heuristic Algorithms for Optimal Nesting of Two-Dimensional Irregularly Shaped Sheet-Metal Blanks .Computers in Industry .1994,24:55—70
- [53] 王英林, 吴慧中 .SOSG 三维布局搜索树的简化研究 .小型微型计算机系统 .1996, 17(2)
- [54] 王英林, 吴慧中 .空间布局求解中的规则匹配 .山东师范大学学报, 1995, 10(3)
- [55] 吴慧中, 王英林 .一种立体空间布局模型及其布局算法 .计算机学报 .1994, 17(11)
- [56] 曹学军, 潘云鹤, 何志均 .基于规划的空间布局 .计算机学报 .1991, 7(7):533—541
- [57] 戴佐, 袁俊良, 查建中, 郭伟 .一种基于八叉树结构表达的三维实体布局启发式方法 .软件学报 .1995, 6(10):629—636
- [58] 戴佐, 查建中 .三维实体布局仿真中干涉检验的八叉树方法 .计算机辅助设计与图形学学报 .1995, 7(4):257—262
- [59] Leung, J .A New Graph-Theoretic Heuristic for Facility Layout . Management Science .1992, 38(4):594—605
- [60] 晏敏, 张昌期, 刘育骥 .基于墙模型的平面布局专家系统框架 . 武汉: 华中理工大学学报 .1990, 18

(1):23—32

- [61] Prasad, S . S ., Dhande, S . G . Computer Aided Design of Optimal Pattern Layouts .CAD/ CAM/ CAE for Industrial Progress, Rajaraman V (ed .), Elsevier Science Publishers B .V .(North-Holland) .1986
- [62] Garijo, F . J ., De Garrido, L . A .A Knowledge Base System for House Design .Proc .IEEE Int . Conf .Syst .Man and Cybernetics .1988:806—809
- [63] Jogn Hunt, Roger Miles .Toward an Intelligent Architectural Design Aid .Expert Systems .1995, 8, 12(3):209—218
- [64] Ramesh Krishnanurti and Rudi Stouffs .Spatial Grammars: Motivation, Comparison and New Results .CAAD Future 's93, Flemming U and Wyk S Van(Editors), 1993 Elsevier Science Publishers B .V .57—74
- [65] 曹学军, 潘云鹤, 何志均 基于规划的空间布局 .计算机学报 .1991, (7):533—541
- [66] 陈国良, 王煦法等 .遗传算法及其应用 .北京:北京人民邮电出版社 .1996, 51—55
- [67] Sriram, D ., Stephanopoulos, G ., Logcher, R ., Gossard, D ., Groleau . N ., Serrano, D . and Navinchandra, D . Knowledge-Based Systems Applications in Engineering Design: Research at MIT .AI Magazine, Fall 1989 ,78—79
- [68] 张钹, 张铃 .问题求解理论及应用 .北京:清华大学出版社, 1990, (12):35—36
- [69] 林峰, 颜永年, 卢清萍, 童水光 基于图形数据的图形参数化方法 .计算机辅助设计与图形学学报 .1993, 5(3)
- [70] 向文, 钱小平等 .基于特征的三维参数化造型 .中国机械工程 .1994, (4)
- [71] 王建业, 王运赣, 张宜生, 陈国清 基于特征的自动建模技术 .华中理工大学学报, 1991, 19(增刊)
- [72] 黄和风, 赵向阳 .构造特征造型 CAD 系统 .中国机械工程 .1992, 3(2)
- [73] Anderson, D . C . and Chang, T . C . Geometric Reasoning in Feature-Based Design and Process Planning .Computer & Graphics, 1990, 14(2): 225—235
- [74] Baud, A .H .and Chang, K .J . Feature-Based Process Planning for Machined Parts . Computer in Engineering, 1988, 1:571—576 Engineering, 1988, 1:557—562
- [75] Ibrahim Ieid .CAD/ CAM Theory and Practive .McGraw-Hill .Inc .1991
- [76] 宁汝新, 徐弘山 .机械制造中的 CAD/ CAM 技术 .北京:北京理工大学出版社 .1991
- [77] 孔庆复 .计算机辅助设计与制造 .哈尔滨:哈尔滨工业大学出版社 .1994
- [78] 唐泽圣, 周嘉玉, 李新友 .计算机图形学基础 .北京:清华大学出版社 .1995
- [79] Wang Jianhua et al, . Sánchez, V . D . A Robustization of a Learning Method for RBF Networks . Neurocomputing, 1995, (9):85—94
- [80] Kung, S . Y . Digital Neural Networks .New Jersey USA: PTR Prentice-Hall, 1993
- [81] 王先逵 等 .CAPP 研究动向浅析 .全国高等学校机械加工自动化研究会论文集 大连:1993, 1—6
- [82] Wang Hsu-Ping, Li Jian-Kang . Computer-Aided Process Planning .Amsterdam : Elsvier Science Publishers B .V ., 1991
- [83] 衷裕水, 范生祥 .CAPP 中回转体零件的全信息描述 .成组技术与生产现代化, 1993, (3): 27—33
- [84] 李德庆, 吴锡英 .计算机辅助制造 .北京:机械工业出版社, 1991
- [85] 赵讷东 等 .基于 DBMS 的工艺计划设计工具 .机械工业自动化 .1995, (1):1—4
- [86] 杨建国, 李培智 .CAPP 开发环境的研究与实现 .机械工业自动化 .1995, (1):5—7
- [87] 陈玉琨, 王宣武, 黄树新 .实用成组技术 .北京:机械工业出版社 ,1994
- [88] 刘宇棱, 易红等 .基于实体造型 CAD 软件加工特征的识别与提取 .成组技术与生产现代化 ,1994

(4):18—22

- [89] 李俊山 . 论工艺设计专家系统初步设计阶段的任务 . 成组技术与生产现代化 . 1993, (4): 22—26
- [90] 蔡力刚等 . CAPP 系统实用工艺数据库与知识库 . 全国高等学校机械加工自动化研究会 . 南京:东南大学出版社, 1994 . 150—154
- [91] 宁汝新, 徐弘山 . 机械制造中的 CAD/ CAM 技术 . 北京:北京理工大学出版社, 1991
- [92] Mital, A . Handbook of Expert Systems Applications in Manufacturing Structures and rules . London:CHAPMAN & HALL, 1994
- [93] 张福学 . 实用传感器手册 . 北京:电子工业出版社, 1988
- [94] 程道喜等 . 传感器的信号处理及接口 . 北京:科学出版社, 1989
- [95] 王正光, 周忠英等 . 数据采集与处理 . 北京:国防工业出版社, 1983
- [96] 魏晓东 . 分散型控制系统 . 上海:上海科学技术出版社, 1991
- [97] 王常力等 . 集散型控制系统的设计与应用 . 北京:清华大学出版社, 1993
- [98] 苏松基 等:最优控制 . 北京:科学出版社, 1993
- [99] 格罗弗, M . P ., 小齐默斯, E . W . 计算机辅助设计与制造 . 北京:科学出版社, 1991
- [100] 沈康主编 . 机械制造工艺及设备 . 北京:机械工业出版社, 1989
- [101] 欧文, A . E . 著, 章兹定译 . 柔性装配系统 . 北京:机械工业出版社, 1991
- [102] 杜祥瑛主编 . 工业机器人及其应用 . 北京:机械工业出版社, 1986
- [103] 安德鲁·库夏克 著 . 杨静宇等译 . 智能制造系统 . 北京:清华大学出版社, 1993
- [104] 国家自然科学基金委员会 . 机械制造科学(冷加工) . 北京:科学出版社, 1994
- [105] 渠川璐主编 . 人工智能、专家系统及智能计算机 . 北京:北京航空航天大学出版社, 1991
- [106] 张伯鹏主编 . 机电智能控制工程 . 北京:机械工业出版社, 1991
- [107] 董春玲等 . 电子精密机械制造工艺学 . 成都:电子科技大学出版社, 1994
- [108] 陈炳森等 . 生产系统学 . 上海:同济大学出版社, 1992
- [109] 张越等 . 机器人化装配过程中的规则与判断支持 . 机器人, 1993, 15(3): 18—22
- [110] 康南生等 . 一个基于知识的智能装配机器人任务级编程系统 . 模式识别与人工智能, 1990, 3(4): 37—45
- [111] 张钹等 . 机械装配规划的自动产生 . 模式识别与人工智能, 1991, 4(1): 11—19
- [112] 曲在滨等 . 智能装配机器人机械零件知识库的自动建立 . 机器人, 1991, 13(2): 25—28
- [113] 杨文源等 . 基于知识 Petri 网的机器人装配规划 . 机器人, 1994, 16(6): 357—361
- [114] 石淼等 . 基于图纸理解的装配关系模型的生成 . 机器人, 1991, 13(5): 23—26
- [115] 侯晓林等 . 面向并行工程的产品装配模型和虚拟制造技术的研究 . 机械科学与技术, 1995(5)
- [116] Homen, L . S . et al . AND/ OR Graph Representation of Assembly Plans . IEEE Trans on Robotics and Automation, 1990, 6(2): 188—199
- [117] Peterson, J . L . Petri Net Theory and the Modeling of Systems . New Jersey: Prentice Hall, 1981 .
- [118] 杨国良主编 . 机械制造 . 上海:上海科学技术出版社, 1992
- [119] 孟兴发等主编 . 机械制造工程概论 . 北京:航空工业出版社, 1992
- [120] Jack, D . Lane . Automated Assembly . Michican : Society of Manufacturing Engineers Publications Development Department, 1986
- [121] Adedeji, B . badiru . Expert Systems Applications in Engineering and Manufacturing . New Jersey: Prentice Hall, 1992
- [122] 潘云鹤 . 中国智能 CAD '94 . 北京:清华大学出版社, 广西科学技术出版社, 1994
- [123] Das, R . and Whiteley, D . The Only Challenging Problems are Deceptive: Global Search by Solving

Order-1 Hyperplanes .Proc .of ICGA,1991,166—173

- [124] Christian Bessiere .Arc-consistency in Dynamic Constraint Satisfaction Problems In Proceedings of AAAI-91 ,Anaheim,CA,1991,221—226
- [125] Pascal Van Hentenryck and Thierry Le Provost .Incremental Search in Constraint Logic Programming .New Generation Computing .1991 (9) : 257—275
- [126] De Fazio, Whitney DE .Simplified Generation of All Mechanical Assembly Sequences .IEEE J of Robotics and Automation .1987 , 3: 640—658
- [127] 李圣怡等 . 智能制造技术基础 . 国防科技大学出版社,1995
- [128] 查建中 . 智能制造工程 . 机械工业出版社,1992