

计算机科学技术发展战略研究

Strategic Research on the Development of Computer S.& T.

中国计算机学会研究组*

一、未来的发展趋势

未来十年计算机科学技术的研究将取得重大进展。主要是:在并行的分布式处理技术、计算机软件、数据库新技术、多媒体技术、科学视算、人工智能和神经网络的基础研究等方面将取得更大突破;计算机电路设计、测试及其工具将得到更广泛的应用;计算机辅助设计、辅助工程、成组技术、计算机辅助工艺编程将得到更迅速的推广;在传统的以磁介质为基础的大容量存储技术不断发展的同时对新的存储介质的研究也将提上日程;计算机安全和可信性的理论和技术将受到更大的重视;汉语言信息处理的研究成果也必然会跨上一个新的台阶。

并行和分布式处理是大幅度提高计算机性能的有效途径。由几台乃至十几台处理机组成的小规模并行机已日趋成熟,应用面也较广。大规模并行计算机尚处于发展初期。目前美国并行计算机的设计和软件已经达到比现有最强大的矢量机高出一个数量级的能力。大规模并行计算机集中于一个单元上,几千台小型处理器紧密地结合在一起,与之相联结的网络又促进了计算机之间的通讯和协同,从而可更有效地利用资源,使其获得比分立计算机更强的解决高度复杂的计算和处理问题的能力。但现有的大规模并行机的应用针对性很强,受到软件的限制,使用也不方便,其峰值速度和实际效率往往有很大差距。这些问题的解决,有待于并行体系结构、并行和分布式算法、并行操作系统、并行语言、并行编译和优化技术等方面的突破,有赖于体系结构、软件和网络一体化研究的进展。

*国家自然科学基金委员会支持的研究项目。组长张效祥(学部委员,计算机学会理事长),本文执笔陈树楷(研究组成员,高级工程师,计算机学会秘书长)。

高效率地生产高质量软件的能力对工业与国防的各种先进技术是至关重要的。现代的高度自动化的系统正变得越来越依赖于软件的可靠性及其功能的先进性,而这些系统的成本也越来越受到软件开发和生产成本的影响。当前,国际上对软件工程方法学和软件开发模型还有争论,如传统的“瀑布模式”和“快速形成原型样机”之争。另一方面,在80年代,大量离散的或部分集成的软件开发工具纷纷出台,几乎覆盖了软件生命周期的各个阶段。在此基础上,计算机辅助软件工程(CASE)工具已有产品问世。它们的推广应用在提高软件生产效率,保证软件质量,降低软件生产和维护成本等方面起到了积极的作用。但是,软件工具离高度集成化和实用化的目标尚有相当差距。以上两个方面,都是综合性强、复杂程度高、规模大的课题,需要综合运用各种软硬件理论和技术持续攻关才能取得预期效果。

80年代后期,面向对象的程序设计(OOP)在沉寂了十多年后又重新掀起热潮。在软件复用技术方面,近年来出现了“软件部件制造商”和“集成软件”的新概念。利用现有成分,可以少量或不用编程,就能组装成新系统的主要成分。程序变换和逆向工程技术也在不断地研究和发展。

关系数据库的理论和技术已相当成熟,但是,现有的数据库没有很好地考虑处理数据语义问题,这一方面的研究还刚在起步。同时,从不同的应用领域提出了许多数据库的非传统应用的课题,导致了一系列数据库新技术的研究,如面向对象的、演绎的、智能的数据库,等等。

通过键盘、鼠标和屏幕显示与操作人员通讯的传统交互方式已远远不能满足计算机蓬勃发展的要求,多媒体技术正应运而生。它是在交互式控制下,把数值、文字、图形、动静图象和声音有机地集成在一起,通过多种媒体使人和机器自然“对话”,使用户有身临其境的“实感”,并把结果综合地“显示”出来。多媒体技术已日臻实用,它的进一

步成熟尚有赖于自然语言理解、声音的识别和合成、图象和声音的更有效的压缩编码和还原技术,以及计算机视觉等基础学科的研究成果。

近年来,计算机图形学在造型技术、画面绘制、人机交互技术、真实感图形显示等方面取得了重要进展,动态三维图形已达实用阶段。科学和工程视算(Visualization in Scientific and Engineering Computing)可利用人类的视觉系统和计算机生成的图象实时地用带有颜色、亮度、彩色动画、正文等的图象生动地把计算过程和计算结果动态地显示在科技工作者面前。在多维数据的算法、科学视算软件、工作载荷的分配以及对计算机网络的需求等方面都有不少课题有待探索。

计算机辅助设计(CAD)被认为是过去25年世界上工程技术十大成就之一,发展迅速,应用广泛,遍及电子、机械、轻工、建筑、艺术等各行各业。它改变了传统的设计方法,提高了工程设计的质量和效益,是计算机应用方面一个极活跃的领域。CAD系统正向着第三代——以工程数据库为核心的集成化系统方向发展。它和计算机辅助工程(CAE)、成组技术(GT)、计算机辅助工艺编程(CAPP)和工厂调度工具(如计算机辅助材料需求计划等)结合起来,成为实现柔性计算机集成制造系统(CIMS)的一个重要部分,将为传统工业的技术改造作出重大贡献。在CAD中,和计算机系统自身有直接关系的是集成电路的计算机辅助设计(IC CAD),除了要研究计算机电路的设计、测试等方面的方法学和工具以外,对某些ASIC电路(专用电路)的设计也应加以重视。

计算机网络是计算机技术与通讯技术相结合的产物。80年代,在局域网和广域网迅速发展的基础上,开放式网络结构已提上日程。在国际标准化组织(ISO)制订的开放系统互连(OSI)标准参考模型上建立起来的开放式网络结构,允许不同厂家的设备方便地互连,具有良好的可扩充性和资源共享能力。与此同时,网络的管理将更趋复杂,网络中数据和信息的安全保密问题将更显突出。

计算机病毒在全世界迅速蔓延,利用计算机行窃、诈骗和破坏的案件不断上升,由客观环境和人为造成的计算机系统事故频频发生,给社会带来了严重的威胁,使信息安全学得到重视和发展。这包括对数据安全、计算机系统(含软、硬件、网络等)安全方面的研究涉及用户识别术、加密术、反窃听术、防复制术及防范和消除病毒的技术等。

与计算机安全有密切联系的另一个分支是计算机的高可信性。尤其是涉及国防、国民经济和社会生活要害部门的联机系统,一旦发生故障,将会造成难以估量的损失。这就推动了容错技术和容错系统的研究。目前美国已提出使计算机在200

年平均时间中连续四年不中断和一次修复只需一分钟的目标。

存储器技术是限制其它信息技术应用的一个因素。高性能计算机应用的开发要求巨大的存储器容量。除了传统的磁介质大容量存储技术有待进一步发展外,光存储介质研究方面的突破也将成为计算机变革的重要支柱。

我国的自然语言处理研究当然是以汉语为主要对象,这就使它成为计算机科学技术领域中一个最具中国特色的分支。汉语不断增长的国际地位和我国科技人员的语言知识结构(至少懂一门外语)都有利于开展汉语和外语的对比研究,更增加了我们在这个分支中的优势地位。语言规律的研究、机器翻译原理与软件环境的研究、面向汉语的文法体系以及语料库的设计和建立都是值得重视的课题。

人工智能和知识工程是计算机科学技术中十分活跃的领域。在知识表示、知识获取、利用知识进行自动推理、自然语言的理解和处理、计算机视觉、机器翻译等方面进行着深入的研究,对机器学习、认知科学和智能机系统的研制等方面进行着探索。专家系统是人工智能和知识工程走向实用化的一个突出例子,在医疗诊断、矿藏勘探、石油开采、军事工程、金融保险、地震和气象预报,以及决策支持等方面发挥着一定作用。但是,现有的专家系统还缺乏足够的知识,各系统的应用范围也极其狭窄,知识库的探索效率还不够理想,有待进一步提高。举世瞩目的日本第五代机计划即将告一段落,取得了一定成果,但离既定目标相距甚远,调子也越来越低。近年来,国际上一度出现研究人工神经网络的热潮,人们试图模拟大脑神经网络来构造新型的智能计算机,这方面争议也很大。归根结蒂,智能计算机的研究和发展离不开人工智能、知识工程、计算机科学、认知和思维科学的理论和实践的支持,在诸多学科取得本质上的突破之前,奢谈“智能计算机”的研制是不现实的。恰好,东方人的传统智慧在研究抽象(逻辑)思维、形象(直感)思维和灵感(顿悟)思维的机制等方面是有优势的。

计算机科学技术的各个分支,都和计算机科学技术理论有不可分割的联系。计算机理论,作为计算机科学的基础,更有一些通用的、根本的和带基础性的课题值得研究,例如:计算模型、算法与对算法复杂性方面的研究等。

二、发展战略研究

计算机科学技术是一门研究内容极其广泛,发展极其迅速,交叉渗透极强,应用领域涉及社会

各方面的科学技术。要全面论述就涉及到学科分支准确划分,这是一件极难且必要性不大的事情。下面仅就国内外的发展情况和预测未来10年可能的动向,分为12个专题分别阐述,提出只属于计算机科学技术中的基础研究和部分应用基础研究的13个重点课题或称优先支持的项目。

1. 并行和分布式处理技术

这是一项综合性的研究,涉及算法、体系结构和系统软件等多方面的内容。并行和分布式算法的研究是基础,并行处理体系结构的研究旨在得到具有高并行处理能力的整机系统,并行操作系统和并行语言编译则致力于充分发挥出整机的效能。该课题包括:

(1)并行和分布式算法。

(2)并行处理体系结构:逻辑增强存储器,计算阵列,超标量(RISC和VLIW等),紧耦合和松耦合等。

(3)并行操作系统的设计和实现技术。

(4)并行语言、编译和优化技术。

2. 新的软件开发模型和开发方法的研究

该项研究主要是指,在一定的软件工程方法学的指导下,对软件开发过程进行抽象,并对该过程中各个步骤进行系统化的描述和规约说明。它强调将开发方法应用于具体的软件开发过程,归纳出系统化的开发模型和软件支撑工具的内部集成模式,在更高的水平上解决软件面向应用的问题。近期强调以面向对象的计算模型和方法为基础,提供有关软件开发过程的规范描述、环境支持和实现手段的技术。

此外,在今后的研究中,还应对各种软件开发方法进行条理化剪裁,研究不同类型的开发方法在软件工程中的应用,进行评价和比较;应设专项研究,以建立各种特定领域中的软件开发模型体系;应将形式化方法的研究应用于软件组合、软件复用和软件信息(数据库、软件库和过程模型支持库)管理等方面。该课题包括:实用的软件工程方法学和规约的研究,软件过程(Software Process)和软件开发模型的研究,面向对象的软件开发方法和技术的研究。

3. 典型的软件工具环境相关技术的研究

该项研究特别强调软件工具集成技术和用户接口技术,这是因为软件工具环境相关技术中的大多数目前仍然单独地、离散地应用于各个宿主系统中,其使用方式主要是手工式的,且缺乏统一的规范。所以,有必要在现有技术的基础上,制定统一的规范和接口标准,大力开发服务于软件工程各阶段的各种工具,并使这些工具配套,即将各项技术进行提炼并融入到统一集成的CASE环境之中。该课题包括:软件集成技术,用户接口技术,

软件复用技术。

4. 数据库新技术的研究

数据库技术对国民经济和科技发展已产生了重大影响。但现有的数据库尚少考虑处理数据语义的问题,也不能满足空间数据、复杂对象的需要,要积极开展数据库新技术的研究。包括:数据模型的研究,面向对象的数据库技术研究,主动数据库系统的研究(含演绎数据库、智能数据库等),联邦数据库系统的研究。

5. 多媒体技术

多媒体技术涉及软件、图形学、人一机接口等分支。包括:图形、图像和声音的压缩编码和还原编码的研究,相关的系统软件和人—机交互接口。

6. 高可信性计算机的结构与评估

高可信性计算机是90年代计算机发展的主要方向之一。目标是:通用机,可用性=99.99%,不可用性=50分钟/年;极高可用性计算机,可用性=99.99999%,不可用性=0.05分钟/年。而80年代,一般运行良好的计算机的可用性=99%,即每周约有100分钟故障停用,离目标相差两个数量级。

亟须研究评估方法才能确认计算机的可靠性、可用性是多少。该课题包括:

(1)多处理器容错系统的结构。采用多处理器是容错系统结构的主流。需要研究系统的组织、存储器的分配与管理、总线安排、操作系统、现场可替换单元。

(2)软件实现容错的算法。用软件实现容错是提高可信性的主要途径。关键是算法,从理论上保证容错。

(3)高可信性系统中的故障诊断。包括芯片内部的内装自测试、电路模块的边界扫描、系统级的在线故障检测和诊断等。

(4)高可靠性和高可用性系统的评估。

7. 信息安全

信息本身就是财富,就是时间。信息不应存储在不设防的计算机系统内,信息的传输不应通过脆弱的公共信道,近年来计算机病毒不断出现,因此研究信息安全刻不容缓。基本上包含:

(1)数据安全。保证信息不被非授权窃取和篡改。密码是可行而且有效的办法。密码作为一种技术源远流长,但作为一门学科,则是近几年的事,特别是70年代中期,数据加密标准DES和公钥密码的提出,它已是十分活跃的一个领域。信息密码已成为计算机科学的不可缺少的内容。

(2)计算机安全。信息存储在计算机中,它的存取也通过计算机进行。过去计算机技术强调对用户的友好界面,但对安全考虑不够。如何保证它的安全,避免事故发生,保证信息不至于泄漏,涉

及到技术和管理等方面问题。

(3) 计算机病毒的防治技术。

8. 标准化的网络体系结构的研究

(1) 没有一个标准的开放式计算机网络系统的体系结构是不能够把各种不同厂家的不同型号的计算机,乃至不同类型的计算机网络彼此互连起来的。国际标准化组织(ISO)提出的开放系统互连(OSI)七层参考模型,就是为了解决这个问题。因而,对于ISO的开放系统互连参考模型(OSI/RM)的研究可以说是计算机网络方面的一项根本的应用基础研究。

ISO OSI/RM分为两大部分:一是低层(包括物理、数据链路、网络和运输等四层),主要与数据通信有关,目前比较成熟;另一是高层(包括会话、表示和应用等三层),主要与系统的应用有关,发展较慢。对于高层,主要是加强其协议本身的研究;对于低层,则研究重点应放在协议验证和一致性测试方面,以使协议的实现建立在可靠的理论基础。当然,对于如何把当前流行的TCP/IP互连网的网络体系结构过渡到ISO OSI/RM这一有很大实用价值的应用研究课题,也应优先支持。

(2) 网络安全和网络管理的研究,也应着重围绕ISO OSI/RM来进行。例如,开展开放系统互连的管理(OSIM)、网络安全体系结构、网络安全服务等研究。

(3) 智能网研究,首先开展智能网的网络体系结构的研究,并把它作为ISO OSI/RM的一种发展来进行。

9. 大规模语言信息处理

大规模语言信息处理研究主要包括语料库设计、机器翻译系统开发、电子词典等内容。基于语言规律归纳的语料库的研究和自然语言文法体系的研究成为大规模语言信息处理的核心,也是其中的难点。

大规模语言信息处理这一研究课题既有实用价值,又有理论价值,它不仅可产生巨大的社会效益和经济效益,而且对于探讨人类理解与习得语言的机制以及人类智能的本质有着特别重要的意义。大规模语言信息处理研究已成为国际上激烈竞争的研究领域之一。欧美日各国纷纷投巨资用于机译系统、电子词典的开发及其相关的语言学理论的研究。为此,建议把大规模语言信息处理作为自然科学基金的重点投资项目给以支持,并以能够处理大规模的真实语料为目标,能切实达到这一目标的理论和方法应是今后支持的重点。具体课题有:语料库的建设与利用技术、适于机器翻译等应用的大规模机器词典开发。

10. 人工智能的基础研究

研究机器智能来模仿、延伸和扩展自然智能

是一件既重要又困难的事情。人工智能的进展与基础研究的突破有着密切的关系。研究的主要内容有:认知模型的研究,自动推理方法的研究,机器学习的研究,分布式人工智能的研究。

11. 人工神经网络

以非线性大规模并行分布处理为主要特征的人工神经网络的研究取得了引人注目的进展。神经计算的非程序的和适应性的信息处理能力将成为新一代计算机的重要基础。主要包括:记忆机制的研究,思维模型的研究,人工神经网络模型的研究,神经计算机体系结构的研究。

12. 大容量信息存储技术

大容量信息存储在近期内应以跟踪国外先进水平为主要内容。应着重支持下列项目:以提高道密度为中心的基础研究已列入1991年度重大项目,磁头—介质界面问题的理论分析,阵列磁盘技术。

各种新的存储材料的研究。

13. 计算机电路的设计、测试及工具研究

(1) ASIC电路设计(PGA, JTAG协议, Fuzzy, Neuron, Bus等);

(2) IC设计中的高层次自动综合,由芯片的自动设计转向系统设计;

(3) Framework设计研究,包括工具集成技术、方法管理、设计过程管理、工具和数据管理,这些都是提高CAD系统效率的关键技术;

(4) 硅编译技术研究;

(5) IC的可测试设计。

(责任编辑 宋义荣)

(上接第19页)

种技术的负面影响;

——调整国际经济技术关系,加强环境生态的全球合作。

上述要点实际上是人们建设生态环境文明,谋求持续发展的基本行动指南。持续发展体现了人类整体的、长远的利益,它强调经济社会发展同生态环境保护必须一致,在一个良好的生态环境中为所有的人提供实现美好生活的机会。无疑这也是一条可持续的发展道路。这条道路是在不危及后人的环境权利,并充分满足他们未来环境资源需求的前提下,满足当代人的发展需求。这条道路不只是在若干年内,对若干地区支持人类进步的道路,也是一条从现在到未来都能支持全球人类进步的道路。这条道路就是以“绿满全球”为标志的持续发展道路。

(责任编辑 蔡德诚)