



虚拟天文台

天文学研究的工具与技术

[英] Matthew J. Graham

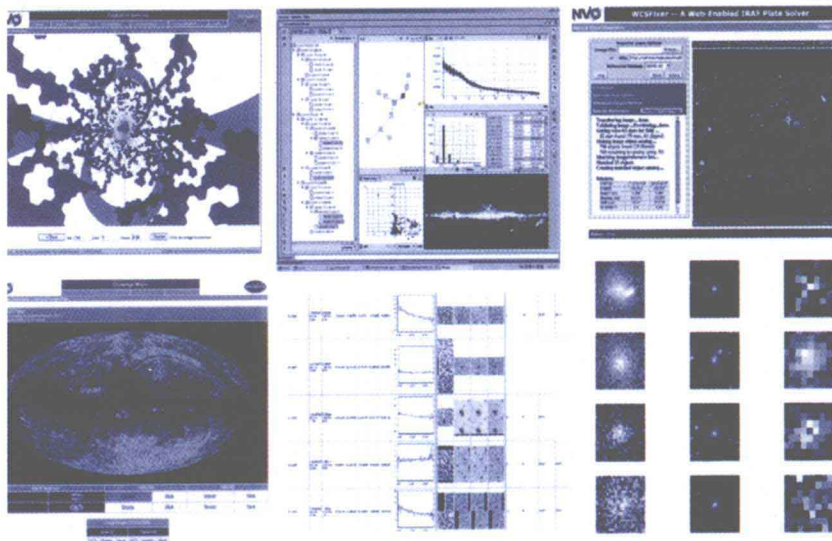
[美] Michael J. Fitzpatrick 编著

[美] Thomas A. McGlynn

崔辰州 等译 赵永恒 校



NATIONAL VIRTUAL OBSERVATORY



中国科学技术出版社

虚拟天文台:天文学研究的工具与技术

[英] Matthew J. Graham

[美] Michael J. Fitzpatrick 编著

[美] Thomas A. McGlynn

崔辰州 等译 赵永恒 校

中国科学技术出版社

· 北 京 ·

图书在版编目(CIP)数据

虚拟天文台:天文学研究的工具与技术/崔辰州等译. —北京:中国科学技术出版社,2010.5
ISBN 978-7-5046-5598-1

I. ①虚… II. ①崔… III. ①虚拟技术-应用-天文台 IV. ①P112

中国版本图书馆CIP数据核字(2010)第062275号

Original Title: THE NATIONAL VIRTUAL OBSERVATORY: TOOLS AND TECHNIQUES FOR ASTRONOMICAL RESEARCH

Copyright © 2008 by Astronomical Society of the Pacific. ASP Conference Series – First Edition

本书中文版由 ASTRONOMICAL SOCIETY OF THE PACIFIC 授权中国科学技术出版社出版,
未经出版社许可不得以任何方式抄袭、复制或节录任何部分

版权所有 侵权必究

著作权合同登记号:01-2009-7136

本社图书贴有防伪标志,未贴为盗版

中国科学技术出版社出版

北京市海淀区中关村南大街16号 邮政编码:100081

电话:010-62173865 传真:010-62179148

<http://www.kjpbooks.com.cn>

科学普及出版社发行部发行

北京长宁印刷公司印刷

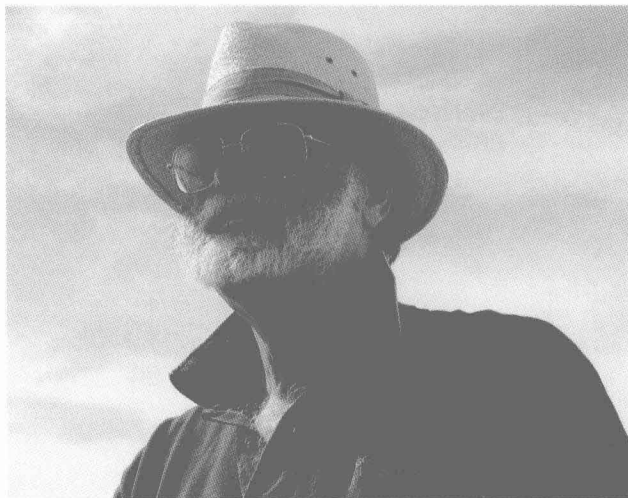
*

开本:787毫米×1092毫米 1/16 印张:38.5 字数:650千字

2010年5月第1版 2010年5月第1次印刷 定价:86.00元

(凡购买本社的图书,如有缺页、倒页、
脱页者,本社发行部负责调换)

敬 献



本书献给我们的朋友和同事吉姆·格雷(Jim Gray)。2007年1月28日,他驾驶着他的40英尺长的“顽强”号快艇,在从旧金山驶向费拉隆岛的途中消失。

吉姆与我们一起工作了很长时间。他给予我们很多的指导,同时给我们带来了灵感,在我们失落气馁之时重新唤起了我们的勇气。他坚信虚拟天文台必将成为未来天文学研究中必不可少的一部分,并与我们一道努力来推动虚拟天文台的发展。在与我们并肩战斗的过程中,他成了一位“地道”的天文学家,为我们的项目贡献了大量的创意和数千行的代码。他的梦想与激情将继续引导我们前进。

Alex Szalay

Bob Hanisch

2007年10月于巴尔的摩

中译本序

虚拟天文台(Virtual Observatory, VO)是通过先进的信息技术将全球范围内的天文研究资源无缝透明地连结在一起形成的数据密集型网络化天文研究和科普教育环境。它将打破现有天文学研究工作的地域与时空界限,通过对资源的统一使用和天文学家的协同工作为天文学带来新的突破,成为21世纪天文学新发现的强大引擎。

20世纪90年代中后期,虚拟天文台的概念首先由美国的天文学家和计算机科学家共同提出。此概念一经提出便迅速得到世界各国天文学界和信息技术领域的高度重视。截至2009年,已有17个国家和地区启动了各自的VO计划。2002年6月,由美国国家虚拟天文台(NVO)、欧洲天体物理虚拟天文台(AVO)、英国虚拟天文台AstroGrid计划共同发起成立了国际虚拟天文台联盟(International Virtual Observatory Alliance, IVOA)。

2002年初,依托大天区面积多目标光纤光谱天文望远镜(简称LAMOST)大科学工程,以中国国家天文台为首的国内天文界提出了中国虚拟天文台(China-VO)研究开发计划。2002年10月,Jim Gray博士访问中国国家天文台。同年,在他的引荐下,China-VO加入国际虚拟天文台联盟。数年来,在中国科学院、国家自然科学基金委、科技部等上级部门的指导和大力支持下,China-VO的研究和建设取得了坚实的进展和丰硕的成果,在天文学和信息技术领域产生了深远影响,在国际虚拟天文台研究领域占据了一席之地。

2008年初,由NVO项目众多骨干成员历时三年多的时间精心编纂的*THE NATIONAL VIRTUAL OBSERVATORY: TOOLS AND TECHNIQUES FOR ASTRONOMICAL RESEARCH*(中文版《虚拟天文台:天文学研究的工具与技术》)一书出版。这是虚拟天文台领域的第一本学术专著。全书共59章,700多页,分7大部分,从科学研究、基础技术、工具应用、系统开发等角度全面地介绍了虚拟天文台的相关知识。

虚拟天文台,和它所凭借的信息技术基础一样,都在快速地发展和成熟着。距离英文原稿的完稿已经有两年多的时间了,书中提及的诸多标准、规范和软件都已经推出了更新的版本,新的技术和系统也在不断涌现。但这些规范和系统背后的思想与理念并没有太大改变,本书仍不失为一部全面了解虚拟天文台的重要参考资料。为了便于国内读者学习和使用,China-VO研发团队决定把这部VO领域的第一本专著翻译成中文。

由于本书内容十分丰富,翻译工作我们采用分工合作的方式进行。前言、丛集著者、软件介绍、序言、第34~36章、第39~44章、第57章、附录3由崔辰州翻译,共15个章节;第1~15章、第59章由杨阳翻译,共16个章节;第16~18章、第45~55章由田海俊翻译,共14个章节;第19章、第25~33章、第37章由刘超翻译,共11个章节;第21章、第56章、第58章、附录1~2由何勃亮翻译,共5个章节;第22~23章由赵青翻译,共2个章节;第38章由吕立强翻译;第20章由崔辰州和何勃亮共同翻译;第24章由赵青和吕立强共同翻译。

为了尽可能准确、流畅地把原著的内容呈现给读者,翻译过程中我们采取了严格的质

量控制流程。基于共同的词汇表和翻译规则,首先分头独立翻译;第二阶段,由译者自己校对两遍;第三阶段,文稿在不同译者间轮换两遍进行交叉校稿;第四阶段,由崔辰州统稿并校对两遍;第五阶段,由赵永恒做最终校稿。整个过程始于2008年春节期间,2010年春节最终完稿,历时整整两年。

衷心感谢中国科学技术大学周又元院士、中国天文学会理事长赵刚研究员、中国科学技术出版社赵晖女士。他们对本译著的出版给予了大力的支持和协助。

由于我们水平有限,书中错误之处在所难免,恳请广大读者谅解并指正。



中国虚拟天文台计划
2010年2月于北京

注:英文原著附有光盘一张,中译本不再提供。有需要的读者请到下面网址下载:

<http://www.china-vo.org/NVObook/>

序 言

虚拟天文台——天文学研究的新环境

Robert J. Hanisch, David De Young

崔辰州 译

简介

虚拟天文台是技术进步和科学实践相互融合的产物,通过对大规模天文数据集提供便捷、高效的访问来推动天文学研究。VO 的起源至少可以追溯到 1988 年,当时 NASA 为其空间科学计划着手建立数据管理中心网络和制订数据保存与再利用政策。20 世纪 90 年代,信息技术的进步为把物理上分布的数据库整合为一个逻辑上统一的系统奠定了基础。天文学家们充分挖掘这些新的计算能力,设计出更大的探测器,无论是从巡天观测还是从定点观测都收获了空前规模的数据。VO 从底层基础到上层建筑为以数据发现、挖掘、融合和统计分析为基础的一种新的科研方式打开了大门。

1. 虚拟天文台的科学动力

数百年来,天文学家习惯了的工作方式是对单个的目标进行独立观测,由不同的人在不同的情况下对同一目标进行重复观测。这种方法对人类发现支配宇宙的一些最基本的天体物理规律起到了很大的推动作用。从牛顿引力定律到恒星内部基本核合成过程的确立,再到广义相对论的验证,一次或者很少几次的观测就足以让我们对身边的宇宙有更深一层的理解。然而,随着天体物理学知识体系的不断扩展以及我们对其理解的不断深化,用一定的物理知识就能很好描述并且通过一组单一的观测就可以解答的重要物理问题越来越少了。当然,这不是说所有的基本问题已经被提出并得到答案,情况远非如此。这里反映的只是问题之间的孤立状态已经变得越来越微妙和难以捉摸这样一个事实。在这样的情况下,现今天文学选择了一个与以往不同却更直接的途径,这就是为每种类型的天体积累越来越大的数据量。

这条途径有两个显著优势。首先,如果有足够的数据,那些只能依靠统计方法才能够解答的天体物理问题便能够得到处理。例如对那些自身观测信号很大程度上被其他效应所掩盖的物理过程的探测就属于这种情况。过去,这样的研究途径即便能够实施也是非

常困难的,因为为了得到统计上有意义的结果不得不把全部的精力花费在收集足够的数据上。此外,因为大型观测设备的观测时间是非常有限的(现在也是如此),许多依赖大量数据才能得以推进的天体物理课题都无法开展。现在,这种途径正经历着巨大而快速的变化。

这种变化是由过去十几年间前所未有的技术进步引发的。为天文学带来这场革命的技术进步主要体现在望远镜设计与制造能力的提升、大规模探测器阵列的发展、计算能力的指数性提高、空间天文台和探测任务工作能力的持续增强,以及覆盖范围和功能都在不断扩展的通讯网络。技术上的巨大进步,包括数字图像(天文学的主要数据来源)以及信息的处理、存储和访问,带来的是可用信息数量和复杂程度的急剧提升。

大规模数据收集的第二个好处是使得偶然性的科学发现成为可能。因此,就抱着一些很普通的天文研究目的而一味的去收集尽可能多的数据也变成了一项有意义的工作。虽然这样的工作会让人感觉很乏味。有人会带着明确的目的对这些数据进行统计分析,而也会有人寻觅那些不期的结果,或者等别人为这些数据提出新的问题。这只是一个简单的例子来向大家表明海量数据广阔的应用前景。

大尺度数字化巡天及其数据产品正在成为天文学数据的主要来源。在这个领域越来越明显地体现出这样一些特征:大尺度的均一化的巡天数据的分析、数百万到数十亿规模上的目标源采样、每个源数十至数百个的测量特性。随着巡天科学日益成为系统化探索宇宙的重要途径,实测天文学的模式正在发生着改变。天空正在经受着跨越整个电磁波段的普查。这样的普查巡天将给我们带来一个全色的、更少偏倚的宇宙图像,至少从理论上说应该是这样。在精心设计的选源标准和观测条件指引下,天文学家已经有能力从不同的空间尺度、时间跨度上在几乎整个电磁波段开展巡天观测。通过生成全色图像甚至宇宙数字电影,以前不可能发现的一些新现象和新规律就可能出现,这些很可能会改变我们对整个宇宙的理解。过去,同一天区内光学图像和射电图像的融合导致了类星体的发现。红外数据的出现带来了被尘埃掩盖的活动星系核和恒星形成区的发现,而这些都是无法从可见光图像中找到的。天空重复性的图像带来了瞬时现象——超新星和更近期的微引力透镜的发现,并对这些时变现象有了更深入的了解。多个大尺度数字巡天联合起来将为参数空间内全新探索的开展提供可能,比如全波段低表面亮度宇宙的研究。

此外,海量数据的存在,在天文学历史上第一次使得复杂的数值模拟结果与完整而系统的多维数据体之间能够进行有科学意义的比较。从速度和地域覆盖上都在快速发展的互联网又让这些科学成果能够方便的为包括美国在内的全世界的天文学家所共享。

虚拟天文台就是基于天文巡天、海量数据和数值模拟来进行整合、比较研究的驱动者。通过为各种数据资源提供一套标准的发现、访问和融合协议,虚拟天文台将推动天文学研究进入一个科学发现的新时代。

2. 虚拟天文台的出现

过去国际上的一些天文项目曾经有着与虚拟天文台相似或相通的目标。其中比较著名的有 20 世纪 80 年代末至 90 年代初受 NASA 资助的天体物理数据系统(Astrophysics Data System, ADS; Murray et al. 1992)、同一时期受欧空局(ESA)资助的欧洲空间信息

系统(European Space Information System, ESIS; Albrecht 1991)、NASA 应用信息系统研究计划(AISRP)1999 到 2000 年间资助的小型探索项目——档案信息访问互联系统(Interoperable Systems for Archival Information Access, ISAIA),以及 NASA 和欧洲的数据中心在 1998 年到 2000 年间联合开发的 Astrobrowse 数据定位服务等。

ADS 和 ESIS 都是在万维网(World Wide Web, WWW)出现之前启动的。虽然项目领导者有着超前的构想,但却缺少搭建基础平台所需的工具。在这些项目提出之后的一个时代里,WWW 逐渐成熟起来。相应的中间件技术,比如 HTTP、HTML、XML、SOAP 等,也得到相当的发展,使得把分布式的查询发送到不同数据中心和服务器的的工作不再需要定制什么额外的软件就可以实现。ADS 在某种程度上获得了成功,尤其是在天文文献索引和交叉引用标准的制定方面。目前,ADS 的文献服务可谓一枝独秀,是国际上使用率最高的天文信息服务系统。它能够在该领域获得成功有以下几个原因:

天文文献的出版商和作者有链接自己资源的强烈愿望和动机。

ADS 设计开发了一个简单而灵活的编码方案,无歧义、易编码、易解码,并且还适于人阅读,这就是“bibcode”。至今这也是该系统的灵魂。

这个编码方案无需替换或者移植什么其他方案。它是全新的并且受到了信息提供者们的欢迎。

ADS 的成功给虚拟天文台提供了明确的经验:迎合天文数据中心参与虚拟天文台的愿望;保持编码标准的简单;基于开放、免费的软件和工具开发 VO 中间件;不要勉强数据提供者改变自身系统。加入 VO 的门槛要尽可能低,收益要尽可能高。

Astrobrowse 项目(McGlynn & White 1998)是由几个天文学家兼技术专家提出的。他们意识到有了 WWW,可以不用从头进行底层技术开发便能够实现 ADS 和 ESIS 项目的基本功能。在没有专门经费支持的情况下,HEASARC、STScI、CDS 的一些人自愿开发了一个数据定位服务,证明了“一站购物”的威力。一条查询语句,比如“告诉我谁有 NGC 1068 的数据?”马上就会导致对数十个数据库系统的访问。

ISAIA 项目(Hanisch 2001)明确了两条重要的设计原则,后来这些都被美国国家虚拟天文台项目所采用。第一条,把摊子铺的过大,在经济上不划算,科学上也没有太大价值。所以 NVO 仅把精力集中在天文学上。当然,NVO 开发的模块也许会很好地在其他领域得到应用,就像 NVO 从别的领域借鉴经验一样。第二条,元数据标准对于数据发现和互操作至关重要。虚拟天文台研究开发过程中,很大精力就要花在元数据定义这项辛苦的工作上。

虚拟天文台(Virtual Observatory)的概念是在美国准备 2000 年新十年调研报告(新千年天文学与天体物理学,(美国)国家研究理事会)的过程中诞生的。理论、计算和数据探索工作组在虚拟天文台概念提出过程中发挥了核心作用。这个工作组认识到天文学正在进入一个大规模巡天和高性能计算与网络能力的新时代,这将为天文学家访问和使用数据的方式带来革命性的变化。最终的报告把(美国)国家虚拟天文台列为优先级最高的小型项目:

委员会推荐了几个地面和空间的小型项目。其中第一个,即国家虚拟天文台(NVO),在委员会建议的小型项目中排在最优先的地位。NVO

将以现在正在产生的大数据库和未来的更大数据库为基础,提供一种“虚拟天空”。它将能够使专业天文学家拥有一种新的研究模式,同时为公众的教育和发现提供一种无可比拟的机会。^①

虚拟天文台实现的前提条件是天文数据有良好的归档管理。在美国,NASA 远见卓识,坚决要求对空间天文项目的数据进行完整和文档齐全的归档,并进一步把这些数据组织成一个分布式的数据系统,让数据库专家和科学家一起工作。一些地基巡天项目也开始学习 NASA 的成功经验,建立全面的、经过仔细定标处理的、文档齐全的数据库系统(比如 2MASS 和 SDSS)。现在,地基天文台也开始采用这种模式。

2001 年 4 月,17 个数据中心、天文台和计算机科学单位自发地联合起来,向 NSF 信息技术研究计划提交了建议书,计划建设 NVO 的基础环境。参与这个建议书的团队在约翰·霍布金斯大学 Alex Szalay 和加州理工学院 Paul Messina(后来由 Roy Williams 接替)的带领下已经一起工作了五年。在这段时间里,项目组成功完成了 VO 基础环境核心部分的设计开发。相关成果已经在应用服务、软件工具和科学范例中得到应用。NVO 正在不断走向成熟,天文界对 VO 能力和潜力的认识在不断提升。越来越多的组织采用 VO 协议共享自己的数据。计算网格技术在 NVO 中得到应用。VO 浪潮还席卷欧洲遍及全球。NVO 与其他 VO 项目一起正在国际虚拟天文台联盟的统领下朝着共同的目标迈进。

3. 虚拟天文台的组成

虚拟天文台天生就是分布式的:数据集仍然属于它们的提供者(天文台、数据中心、巡天项目、研究团组),通过标准的接口进行访问。图 1 显示了各种组件和接口之间的关系。

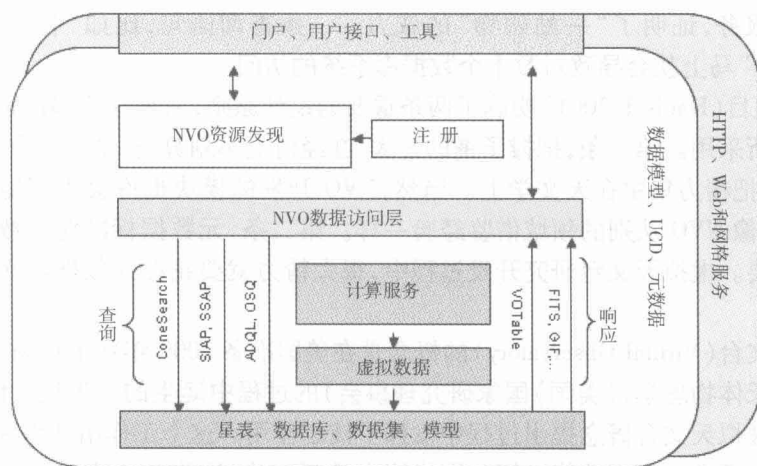


图 1 虚拟天文台的总体架构

① 新千年天文学和天体物理学,(美国)国家研究理事会,(美国)国家学术出版社,2001,p. 14.

对于用户来说,访问 VO 的入口是科学应用程序或者工具——门户,它们利用 VO 协议发现和访问数据。通常情况下,这些门户通过查询注册(registry)来获得 VO 内数据资源的信息。注册系统收录了能够通过 VO 接口进行访问的数据集、数据库、星表、巡天、计算服务等的信息。

然后,具体的图像、光谱、星表、数据库便通过数据访问层(DAL)一系列标准的协议进行查询和访问:

- 使用锥形检索(Cone Search)对星表进行基于位置的搜索
- 使用简单图像访问协议(Simple Image Access Protocol,SIAP)搜索并获取图像
- 使用简单光谱访问协议(Simple Spectrum Access Protocol,SSAP)搜索并获取光谱
- 使用天文数据查询语言(Astronomy Data Query Language,ADQL)通过 Open SkyQuery(OSQ)门户查询数据库

有些数据请求要求进行实时的数据格式转换或者重新采样。这些数据在被访问之前是不存在的。这些虚拟数据由数据提供单位或者高性能计算中心的计算服务动态生成。数据通过标准的格式发送给用户,比如 FITS。图像可以使用 GIF 或者 JPEG,表列数据可以使用 VO 的标准格式 VOTable。数据访问协议是在标准的数据模型基础上设计的,并使用标准的元数据来描述。统一内容描述符(Unified Content Descriptors,UCD)就是其中的一种,它用来明确地标识星表和数据库中列的内容。这些协议和数据格式标准将在本书后面的章节中进行详细介绍。

4. 虚拟天文台与其科学梦想

VO 的概念在全球已经发展了五年多的时间。不同形式的实施方案正在接近初步成熟。基础架构基本到位,一些基本的工具和应用服务被开发出来并得到验证。因此,正如前文所述,VO,特别是 NVO,已经可以为天文学家所用。当然,相关的工具和应用服务还有很大需要完善的空间。

为了满足不断发展的超大型数据集对 VO 的需求,我们和国际上的 VO 伙伴一起开展了许多工作。这些工作大多建立在以指数性增长的计算、软件和网络能力基础上。VO 这种与空前复杂和空前投入的新观测设备的结合很容易给人一种印象,就是 VO 是“大科学”的产物,是为“大科学”服务的工具。如果真是这种情况,那我们会毫不犹豫地断言虚拟天文台的想法失败了。毫无疑问,把当今和未来各式各样的数据集,无论大的还是小的,连接融合在一起是 NVO 和类似项目的目标。同样不可否认,VO 正力图把每个巡天的成果都纳入其中,会充分利用不断发展的计算机网络技术。还有,NVO 和其他项目也肯定会为大科学和大用户的需求服务。但是,VO 理念下一个最基本的观念是让天文界和教育界的所有用户都能够方便地使用天文数据和工具。为了最后的成功,NVO 和它的国际合作伙伴们必须完成这个任务。

因此,VO 的终极目标是能够成为“人人”的 VO,成为所有对天文和天体物理有兴趣的人的日常工具。当 VO 成为一种想当然的事物,当它的存在已经不再被人提起,这时候它才真正成为了天文学基础架构的一部分。当把支持强大研究的工具返还给一个个天文学家的时候,不管他们是哪个单位哪个组织的,VO 推动天文学研究的目标才得到了圆满

的实现。此时,VO 让所有天文学家都可以开展丰硕的研究,唯一需要的就是产生好想法的洞察力。因此,虚拟天文台将会成为推动 21 世纪天体物理研究进入新天地的一架马车,为大科学、小科学提供平等的机会,为全世界的研究团队和每个天文学家提供同样简单的资源访问,还有前所未有的速度、深度和广度。

致谢

本节关于虚拟天文台的介绍,内容源自过去五六年里 NVO 开发项目成员的报告和讲稿,现在已经不可能明确指出哪部分来自哪个作者。除了本章的作者,我们在此特别感谢 A. Szalay、R. D. Williams 和 S. G. Djorgovski。

我们感激 NVO 开发团队和国际虚拟天文台联盟的合作者们所作出的贡献和创新。NVO 是一次奇妙的科学与技术发现之旅,研究如何跨组织和国界开展工作,力争把一种全新的数据组织、整合和理解方式变成现实。我们感谢为 VO 事业作出贡献的所有的人。

参考文献

- Albrecht, M. A. 1991, in "Databases and On - Line Data in Astronomy," ed. M. A. Albrecht & D. Egret, Astrophysics & Space Science Library 171, 127
- Hanisch, R. J. 2001, in "Virtual Observatories of the Future," ed. R. J. Brunner, S. G. Djorgovski, & A. Szalay, ASP Conf. Ser. 225, 130
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C. S., Murray, S. S., & Watson, J. M. 2000, A&AS, 143, 41
- McGlynn, T., & White, N. 1998, in "Astronomical Data Analysis Software and Systems VII," ASP Conf. Ser. 145, 481
- Murray, S. S., Brugel, E. W., Eichhorn, G., Farris, A., Good, J. C., Kurtz, M. J., Nousek, J. A., & Stoner, J. L. 1992, in "Astronomy from Large Databases II," ed. A. Heck & F. Murtagh, ESO Conf. Proc., 387

相关链接

美国国家虚拟天文台 (NVO), 2006 年 11 月, <http://us-vo.org/> [2006 年 11 月 30 日访问]

国际虚拟天文台联盟 (IVOA), 2006 年 11 月, <http://ivoa.net/> [2006 年 11 月 30 日访问]

前言

在过去的三年里,每年9月初都会有一群人聚集在科罗拉多的阿斯潘(Aspen),来一起学习虚拟天文台知识,规模大约在40人上下。从世界的各个角落来到这里,这群人的身份可谓是五花八门,从年轻的研究生到天文学教授,从计算机新手到软件工程专家。通过学术报告和手把手的指导,在一个多星期的时间里,他们学习什么是VO,如何使用它。在专家的启迪和协助下,他们会利用一天半的时间设计一些VO的新应用。也许只是简单地利用VO的协议对天文学家发布一个数据集,或者把一段已有的代码包装成一个VO服务;也可能是通过把现有的VO数据和服务综合起来而取得的一项令人兴奋的科学新结果;甚至还可能是为VO自身带来的革命性的贡献。

对于这几期NVO暑期培训班(NVOSS),学生们给予了非常高的评价。于是,在2006年底我们决定2007年不再简单地组织一个会议,而是更进一步,要把过去几年的经验落实到字面上,与更多的人一起分享。许多教员、学生以及NVO项目的成员都对本书作出了贡献。大家共同把那种在Aspen色彩绚丽的秋天、人与野狗熊和平相处、群山环抱的惬意环境中学习天文新技术的精髓提取出来。希望本书中提供的这些资料能够促使您在日常工作中使用VO,更希望您能给VO带来新颖别致的应用。

1. 本书读者对象

虚拟天文台是天文学和计算科学相结合的产物,因此本书的潜在读者群跨度很大,从仅仅知道自己所属星座的软件开发人员一直到连收发Email都不熟练的天文学教授。不过由于书中大部分的内容是与计算程序相关的,因此本书理想的读者很可能是具有上述两种背景的结合者,比如对天文学感兴趣的开发人员或者是有一些编程经验的天文学家。不管怎样,作为本书的读者,我们希望您至少能使用网络浏览器。

2. 如何使用本书

基于不同的兴趣、知识背景和使用目的,天文学家和开发人员将会从不同的角度接触虚拟天文台。书中的章节大多是相对独立的,读者可以依据自己的兴趣进行跳跃式的阅读。当然如果愿意,您也可以从头到尾通读本书。请根据您所代表的用户类型和预期的目标来选择自己的阅读方式。

简介部分介绍了VO的背景,描述了VO产生的科学动机,对VO的各个组成部分进行了简要介绍,还讨论了要把VO变成现实所面临的挑战。这个部分还对本书附带光盘中的软件进行了介绍。这些软件大部分是在NVO暑期培训班上作为教程的一部分使用或者开发的。并不是每个读者都需要安装这些软件,但软件安装过程对那些希望随本书的介绍进行相关工具练习的朋友来说肯定是有帮助的。

本书的其余部分按照章节共有的主题分成了几个大的部分,比如“工具”、“数据发

布”等。当你以某个主题为线索对本书进行学习时,比如数据发现、VO 协议、客户端应用、开发工具等,都可能需要跨越多个部分。关于书中各部分及其可能的用途在表 1 中进行了概括。

表 1 本书章节概况

章节	标题	描述
I	VO 工具介绍	当前可用的一些高层 VO 工具。许多在附带的光盘中已经包括。
II	虚拟天文台的科学	利用上述工具能够完成的一些科学工作实例。还讨论了如何将 VO 与天文学家目前使用的一些环境整合在一起。
III	技术与协议	VO 系统中采用的一些关键格式和技术。这些基础知识有助于了解 VO 是如何工作的。
IV	VO 环境下发现和注册资源	通过注册来在 VO 中寻找和发布数据的基本知识。
V	VO 中的数据访问	向数据用户和发布者介绍数据访问相关协议。
VI	在 VO 中发布数据	在 VO 中创建新的数据服务。
VII	基础技术	SQL、XML、Web 服务技术的简单介绍,许多 VO 协议和服务的实现就是建立在这些技术的基础上。

2.1 推荐的阅读方式

首先,读者可以判断一下自己属于下面描述的几类读者中的哪一种。本书的自然顺序是期望让一个 VO 新手通过阅读这些资料便能够完成一个典型的研究计划,很多情况下这些计划与过去参加 NVOSS 的学生所做的计划类似。虽然本书不是特意做成一本参考资料,但由于每章的内容基本都是独立成篇的,读者可以依据自己的兴趣挑选来读。

有兴趣尝试和了解 VO 的科学家们:建议这些读者从第二部分的头几章开始,这部分介绍了在 NVO 暑期培训班上大家用 VO 完成的一些科研课题。这些 VO 新手便可以使用本书第一部分介绍的一些网页式的工具,来开始自己的 VO 科学探索体验。具体说来可以从下列几步开始:

- 利用 DataScope 和 Registry 发现数据(见第五章和第七章)。
- 利用 WCSFixer 修正图像坐标(见第十六章)。
- 利用 Open SkyQuery 完成星表间的交叉相关(见第十三章)。

中级读者也许已经安装了附带光盘中的软件,并且还希望下载更多的服务和数据。当他们试图把 VO 数据融入到自己的分析环境中时就会发现书中第一部分介绍的一些单机程序会在数据操作、访问和分析等方面为自己提供很大方便。例如:

- 可以利用 SkyView 来抽取图像(见第五章)。
- 利用 TOPCAT 和 STILTS 进行交叉认证和表的转换(见第二章和第二十一章)。
- 在进行数据分析时用 Aladin 对图像和星表进行合并(见第四章)。
- 还可能会用到本书中介绍的其他应用和程序。

第三部分介绍了所有 VO 服务共享的一些技术:用来进行高级数据查询的天文数据查询语言(ADQL),作为 VO 喉舌的 VOTable 数据格式标准,在注册和数据服务中使用的数据模型和元数据标准。第四部分简单介绍了在 VO 注册中如何发现已发布数据和服务的内容,以及如何利用标准的 VO 协议去直接访问它们。第五部分对这些协议进行了更进一步的介绍,第六部分则重点阐述了如何利用读者自己的客户程序去访问数据。

当科学家们对 VO 的能力及其数据资源有了一定程度的了解后,他们可能就会希望把这些与自己的研究更紧密地结合起来。第二部分中关于把 VO 服务整合进天文台业务的论述(见第三十一章)从研究所的层面上介绍了如何实现上述目标。对于个人研究而言,对那些自己开发分析处理程序,希望把 VO 功能整合到已有的比如 IRAF 或者 IDL 环境中的天文学家,可以在第二部分的结尾找到相关的讨论。此外,在第一部分的最后对其他一些编程接口作了介绍,包括 VOClient(C、Fortran 等许多语言)以及 Python、IDL 两种语言的 VOLib 工具集。对那些计划在其他环境中使用 VO 的用户,这些章节仍然是有参考价值的,它阐述了 VO 如何以自己的能力来补充现有研究环境的功能。

数据发布者:VO 并不是仅为那些希望访问别人数据的人准备的。科学家和研究机构也希望向 VO 发布特定资源的数据和信息。第四部分介绍了 VO 注册以及检索注册的客户程序和方法,同时还介绍了数据发布的基本知识。这部分对于初次发布数据的人会有很大帮助,能够从这里了解到正确注册自己的数据都需要什么,同时还能知道别人将会如何发现自己的数据。

数据发布者们还需要对第五部分中讨论的发布数据时将要用到的一些协议有所了解。如果希望发布表列数据,那么锥形检索服务或者 SkyNode 服务是合适的选择。如果发布的是图像,则选择的应该是简单图像访问(Simple Image Access)服务。光谱的发布则需要使用简单光谱访问协议(Simple Spectral Access Protocol)。继第五部分对这些协议进行了理论上的论述后,第六部分具体地介绍了如何搭建新服务,以具体的例子讲述如何使用这些基本协议来搭建简单服务。

要把数据共享出去,除了需要构建一个软件系统外,发布者们还必须理解他们需要怎样描述自己的数据。第三部分中的章节就描述了基本的元数据,介绍了如何表述数据覆盖和位置,以及如何参与有关实时事件的网络信息交换。可以翻阅第三部分关于 VOTable 的章节,以保证自己写数据的时候是按照 VO 数据交换的标准格式进行的。

应用和服务开发人员:对于那些计划开发新的应用程序或者改造已有程序的开发者们来说,VO 提供了一系列程序库来实现无论是从底层还是上层与 VO 的完全整合。第四十二章和第四十三章讨论了如何查询 VO 注册(第四部分)。第五十章介绍了如何在不同的编程语言环境下(第一部分和第五部分)利用 VOClient 库通过锥形检索,SIAP、SSAP 协议获取数据。第二十三章和第二十四章介绍了在 Python 和 IDL 环境中如何利用 VOAgent 和 VOLib(第一部分)实现类似的功能。在第二部分最后深入讨论了如何把 VO 的功能整合到常用的分析环境中以及如何改造现有的应用程序。第五部分关于协议的详细介绍阐述了开发者如何设计个性化的服务。如果用户希望开发新的 VO 服务或者应用程序,我们建议他从研究现在已经发布的这些软件库着手。

第三部分的内容对开发者们了解 VO 各种功能的内在工作方式很有帮助。尽管这些

章节没有讨论各种协议和服务的细节,但描述了一些底层的标准。VOTable 格式在数据协议中无处不在。任何需要了解数据访问内在技术的用户都必须先理解 VOTable。同样,ADQL 被用于查询 SkyNode 和注册。比简单的位置和关键词匹配查询更复杂的检索语句,只要目标服务支持,ADQL 将是最佳的解决方案。

本书第七部分对所有这些技术进行了非常简略的介绍。这部分主要是让具有不同背景的读者对这些底层技术有所了解,以便更好地理解这些技术如何在 VO 中得到应用。如果想深入学习 SQL 查询语言和 XML,可以参考其他很多相关资料。

3. 本书未作介绍的内容

当前,VO 的各个方面都在不断发展之中。无论是本书还是其他的类似工作都只能是时下状况的一个局部的快照。VO 许多方面的成果我们都没能在本书中容纳进去,比如:

- PLASTIC (<http://plastic.sourceforge.net>),一个允许客户端应用间通讯的协议,以及 IVOA 应用工作组在应用消息方面更多的工作。

Astrogrid Workbench 和 Common Execution Architecture(公用执行架构,CEA),用户可以利用它们构建并执行工作流。

- IVOA 安全框架,其中包括单点登录(SSO),委托与群体授权服务。
- IVOA 语义工作组的工作以及其他本体方面的内容,还有语义网技术在天文学方面的应用。

如果想深入了解上述内容可以访问 IVOA 网站(www.ivoa.net)的相关部分。我们也希望在本书今后的版本中能够补充这些内容。

4. 本书约定

本书中用到的印刷体例包括:

斜体 (*Italic*)

用以表示某个概念的第一次出现。

定宽字体 (Constant width)

表示计算机输出或者程序代码。

> 加黑的定宽字体

表示用户输入。

5. 致谢

我们特别感谢 Sarah Emery Bunn,对她高超的组织和管理能力表示由衷地钦佩。

通过美国自然科学基金会 (NSF) 与约翰·霍布金斯大学达成的合作协议 (编号 AST0122449),美国国家虚拟天文台项目得到美国自然科学基金会的资助。书中的大部分材料最初是为 NVO 暑期培训班准备的。NVO 暑期培训班得到了美国自然科学基金会

和美国宇航局的联合资助。我们对他们的支持表示感谢！

查善隼

编著

Matthew J. Graham

美国加州理工学院

Michael J. Fitzpatrick

美国国立光学天文台 (NOAO)

Thomas A. McGlynn

戈达德太空飞行中心 (GSFC)