

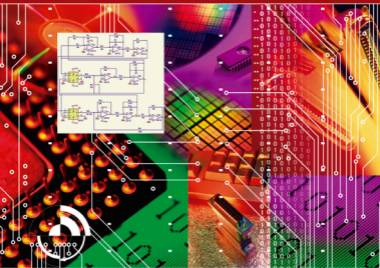
2013年第1期(总第10期)

简讯

国家数学与交叉科学中心

National Center for Mathematics and Interdisciplinary Sciences, CAS

<http://www.ncmis.cas.cn/>



《简讯》编委会 2013年1月

科研进展

量子控制研究取得进展

文:信息技术部

目前信息技术研究部的齐波等在量子调控的基础理论研究方面取得进展。

量子力学理论是上世纪最重要的科学发现之一,它极大的改变了人们对自然界的认识。量子力学的发展始终伴有新奇物理现象的发现,对这些现象的不断研究不仅导致大量新概念的建立和基础理论的变革,而且引发了新的科学技术革命,使得量子调控已经成为可能,甚至成为基础理论和技术进一步发展的必需。理论与实验研究表明,新兴量子技术可以完成许多经典技术几乎无法完成的任务,不仅对国防建设和国防安全至关重要,而且可以极大的改善人们的生活方式。新兴量子技术的发展尤其是其规模性实现需要系统的控制理论为其提供必要的支撑。

相较于经典系统,量子系统具有许多独有的新奇的特性。比如测量量子系统会导致量子系统状态的随机坍缩;著名的海森堡不确定性原理更是从本质上限定了量子系统测量能够得到的信息。我们课题组紧紧围绕量子系统独有的特性以及控制理论的核心概念开展了系列研究。

首先研究了量子反馈控制与开环控制的特点与极限能力。正如如上所述,对量子系统进行测量在获得量子信息的同时,不可避免地造成量子态的随机坍缩,导致我们必须在控制过程中对付该量子系统所固有的测量坍缩效应,增加了量子调控的困难。我们针对一类比较普适的控制模型,基于量子态的鲁棒制备这一基本问题,1)首次给

出了测量通道选取的充要条件;2)证明了在对付初态不确定性方面量子反馈控制的控制效果仍然优于开环控制;3)进一步给出了在测量通道与系统自由哈密顿量不对易情形下,采用量子反馈控制来制备系统自由哈密顿量的特征态时的极限能力。相关结果已经被 IEEE Transactions on Automatic Control 接收。

进而我们又探索了如何基于实际量子系统而非基于其模型设计控制策略。由于任何模型只是实际系统的一个近似,因此基于模型设计的控制策略即使可以证明其施加在模型上时有很好的控制效果,但是当其真正施加到实际量子系统上时控制效果却可能非常不理想。我们将实际系统与其模型的差别视作在模型的基础上附加的不确定性。针对一类比较普适的系统,基于量子态的鲁棒镇定这一重要的控制任务,结合量子系统区别于经典控制系统的独有特点,提出了一个两步分离原理来针对实际量子系统设计控制策略。具体地,第一步首先根据已有精确信息建立实际量子系统的模型,然后根据模型设计反馈控制策略;第二步将实际系统与模型之间的区别视为可以导致退相干的噪声,利用耦合量子系统的张量积结构,运用动态解耦技术实现第一步中所设计的反馈控制策略。我们证明了在适当的条件下,该分离原理的有效性。相关结果已经被 Automatica 接收。

高档数控系统的最优插补方面取得进展

文：先进制造部

高档数控研究课题组在高档数控系统的最优插补方面取得一系列进展。

高档数控机床是国家打造制造业核心竞争力的战略性装备,数控系统是数控机床的“大脑”,是决定其性能的关键因素。目前高档数控系统的技术发展趋势是高速、高精度、高效率。数控系统的若干核心技术,如最优插补、动力学分析与误差补偿,是实现高速、高精控制的基础。目前我们研究的重点是数控系统的最优插补。

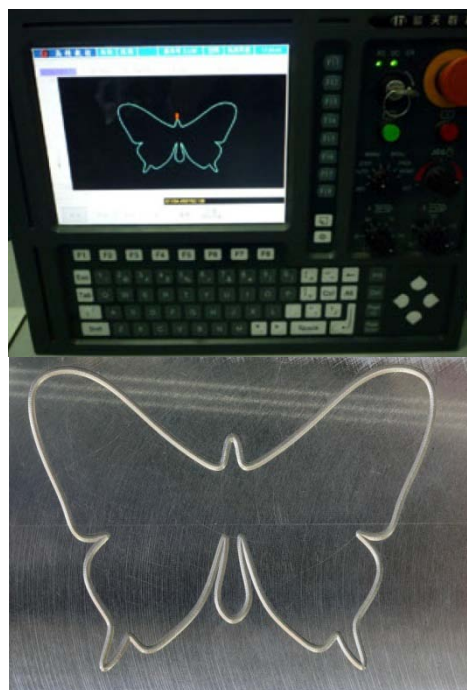
高速高精数控的要求需要我们考虑机床的基本性能与加工过程中的动力学性能与几何变化的情况。数控系统所采用的速度规划算法,将直接影响数控系统的运动精度和运行效率。针对这些问题,我们提出了几种新的方法。

高速高精数控就是要求我们在快速加工工件的同时,能够高质量地完成数控加工。高速加工,就是要求我们在给定的约束条件下设计时间最优的加工方法。而在引入高速度的同时,我们必须要考虑高精度的要求也就是加工误差,而引起加工误差的一个重要因素是弓高误差。这方面主要包含以下两个工作。

1、由于时间最优插补方法针对一般曲线(三次以上样条或 Nurbs 样条曲线)非常复杂,我们针对三轴机床,提出了离散化的方法,给出了在机床速度,各轴加速度下的时间最优的速度规划方法。进一步地,由于加速度的突变会导致机床震动,从而影响加工质量,我们通过线性规划的方法,给出了速度光滑化的方法,从而使得最终的加速度曲线除了极个别点外都是连续的,而这

一过程的所损失的时间在可接受的范围之内。仿真实验与实际加工表明,我们的算法是高效且稳定的,见下图(数控系统与加工实例)。

2、我们给出了机床加工速度,切向加速度与弓高误差限制下的时间最优插补算法,这一方法将弓高误差转化为速度限制曲线,从而可以全局保证弓高误差的限制。我们与中科院沈阳计算



所合作,设计了新的 G 代码,并在机床上实现了这一算法,进行了实物加工。

另外,从减少震动的角度来看,我们必须设计专门的加减速控制规则。而高速高精数控系统中,更应避免加减速结束时的加速度突变,以减小机械冲击,得到好的加工效果。实际测验告诉我们,对 Jerk(速度对时间的一阶导数)或者进一步地,对 Jounce(速度对时间的二阶导数)进行限制可以有效地减少震动,从而提高加工质量。在

这方面,主要有以下几个工作:

1、在给定机床加工速度、各轴加速度以及各轴 Jerk 约束的前提下,我们给出了给定路径下的时间最优的贪心算法。我们证明了时间最优的速度曲线必然是“Bang-bang-singular”的 即在任一时刻,上述约束总有一个达到边界值。通过给出关于速度的二阶微分方程的闭形式解,我们给出了最终速度规划曲线的具体表达式。仿真实验表明,我们的算法是可行的。

2、针对五轴数控系统,由于上述方法求得的曲线表达式较为复杂,而且 Jerk 约束的表达式是非线性的,针对这一问题,我们设计了离散化的方法,通过将 Jerk 约束适当放缩,将其转化为一个线性约束,从而将速度规划的问题转化为线性规划问题,其复杂度是 $O(N^{3.5})$,其中 N 为一段曲线上离散点的个数。仿真结果表明我们的算法具有很高的实用性。

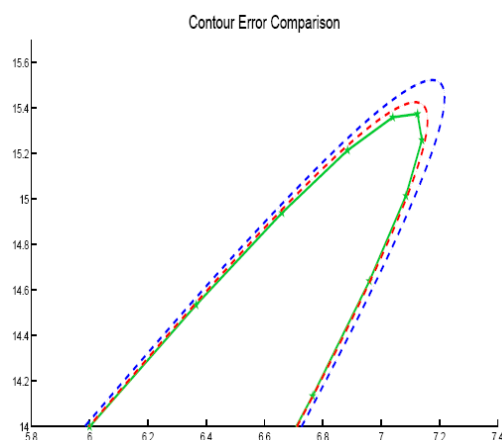
3、在给定弓高误差限制、机床加工速度、切向加速度、切向 Jerk 与切向 Jounce 约束下,我们给出了时间最优的加减速模式以及它们之间的转化关系。提出一种基于样条曲线的加加速度连续的速度规划算法,避免了加速度的跳跃。速度规划采用“七段”式的 Jerk 变化方式,既做到 Jerk 的连续,又尽量加快加工速度。最后,将所提出的算法应用到实际的加工仿真实例中。从最终的仿真结果与实际加工来看,基本上达到了预期的效果。

4、对于五轴数控机床,由于两个角度的转动会引起平动,因此,其速度规划算法会非常复杂。由于上述几个方法给出的约束并不在一个坐标系下,例如,关于机床加工能力的速度、加速度、Jerk 有界是在机床坐标系下的约束,弓高误差是在工件坐标系下的约束。针对上述问题,我

们提出了离散化的方法,将上述约束下的速度规划问题转化为非线性规划问题,给出了沿给定路径的速度时间最优速度规划算法。我们同样证明了在上述约束下的时间最优速度规划满足“Bang-bang-singular”条件,即在任一时刻,上述边界限制总有一个达到。从最终的仿真结果上看,基本上达到了预期的效果。

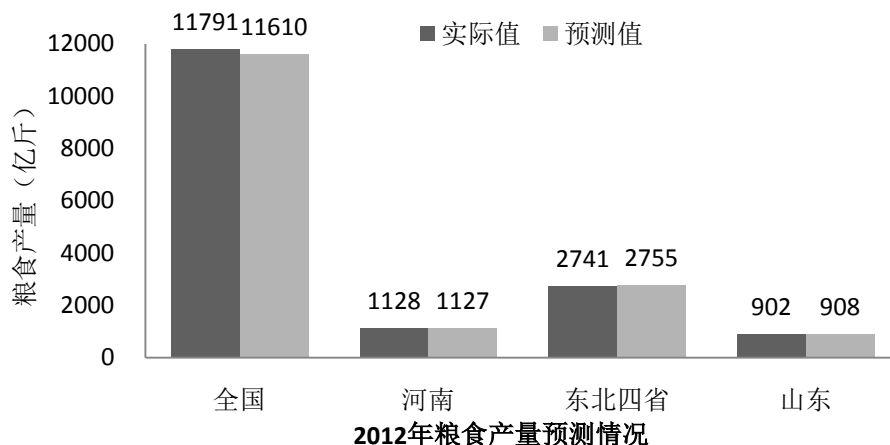
5、此外,由于机床的运动并不是严格按照人们设想的轨迹运动,根据不同的控制器,理想位置与实际机床位置有所偏差,这一偏差称为是随动误差。我们针对一类 PD 控制器,研究了在速度、分轴加速度以及随动误差约束下的最优插补问题。由于随动误差约束是非线性的,我们通过对方程进行简化,可将这一问题转化为一个凸规划问题,但即使如此,问题求解仍然非常复杂。进一步的,我们通过对随动误差的不等式进行适当地放缩,将这一问题转化为线性规划问题,从而给出了离散点数 N 的 $O(N^{3.5})$ 的复杂度算法。仿真实验表明,采用这一方法,可以得到更精确的加工曲线。

具体见下图(绿色实线为待加工轨迹,红色虚线为加入随动误差约束后的加工轨迹,蓝色虚线为未加入随动误差约束时的实际加工轨迹)。



2012年粮食产量预测达到很高精度

文：经济金融部



经济金融研究部陈锡康、杨翠红课题组在2012年我国粮食产量预测工作中再创佳绩。与国家统计局发布的统计数据相比,课题组做出的全国以及主产省区的2012年粮食产量预测结果均达到很高精度。

2012年4月26日,中心经济金融研究部陈锡康、杨翠红研究员等完成了“2012年全国粮食、棉花和油料产量预测报告”。报告预测,今年我国粮食将获得丰收,预测全年粮食产量约为11610亿斤,比上年增产186亿斤左右。国家统计局于2012年11月30日发布公告,2012年全国粮食总产量为58957万吨,比2011年增产1836万吨,即11791.4亿斤,增产367.2亿斤。课题组关于我国粮食产量将实现“九连增”的预测得到证实。全国粮食产量预测值与实际值相比,误差仅为-1.54%。同时,上报中央的预测报告得到了多位国家领导人的重要批示。

2012年5月份课题组也陆续上报了河南、东北四省区、山东等主产省区的粮食产量预测报告。预测2012年河南省、东北四省区、山东省

的粮食产量将分别达到5635万吨(1127亿斤)、13775万吨(2755亿斤)和4538万吨(908亿斤)。实际结果显示,河南、东北四省区和山东粮食产量分别为5638.6万吨(1128亿斤)、13703.5万吨(2741亿斤)和4511.4万吨(902亿斤)。预测结果与实际产量相比,误差分别为-0.06%、0.52%和0.59%,达到了非常高的预测精度。上报省区政府部门的预测报告得到了河南省副省长刘满仓、河南省发展和改革委员会主任张维宁等领导的重要批示。例如,张维宁主任认为“.....这将对安排我省秋粮生产、消费和调入调出等起到重要的参考意见”。

粮食安全是国家政治、经济、社会安全的重要组成部分。中央及地方政府一致高度关注粮食生产。提前期较长且预测精度较高的粮食产量预测结果可以为我国及时安排粮食进出口以及地区间粮食的调入调出提供参考。中心经济金融研究部每年撰写的高质量粮食产量预测报告为中央及地方政府有关部门科学决策提供了重要的参考依据。

地球冰盖模拟取得进展

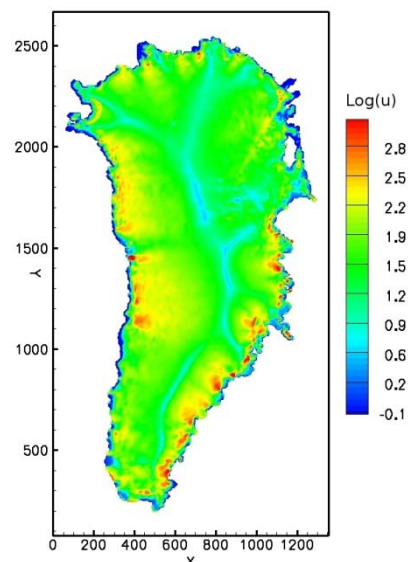
文:材料环境部

冷伟等近日在地球冰盖模拟方面取得进展。他们针对冰盖模拟的完全模型,提出了高效的具有可扩展性的求解方法,开发了相应的并行计算程序并进行了模拟,获得了与以往模拟相比跟实际观测更吻合的结果。

全球气候变化会对人类活动造成重要的影响,地球温度的升高使得冰盖消融,引起海平面上升,导致人口密集的沿海地区面临咸潮破坏,甚至淹没之灾。冰盖模拟的内容为在一定的的大气条件下冰盖的移动、融化和断裂。预测出这些数据,就可以得到由冰盖进入海洋的总水量,然后计算出冰盖对海平面上升的影响。由于冰盖问题的计算量很大,所以目前冰层的研究中存在着很多近似模型,但这些模型对完全的模型进行了不同程度的简化,因此也带来了很大误差,特别是在人们感兴趣的冰层边界,这些误差可能非常大。随着近年来计算机运算能力的大幅提高,并行计算技术的不断完善,以及大规模并行计算的机群在逐渐兴建,如何利用大型机群的超强计算能力,发展冰盖问题的非线性的完全模型的高效求解方法,进行地球冰盖模拟已成为冰盖研究热点。

他们在地球冰盖模拟方面取得重要进展包括:对冰盖的完全模型,推导了非线性问题的 Newton 迭代格式,相对于以往文献中普遍采用的 Picard 迭代,大大减少了求解非线性问题的迭代次数;对线性化的模型导出的鞍点问题构造了新型预条件子,该预条件子从算子运算的角度出发是对原问题的很好的逼近,在该预条件子中,又采用了代数多重网格方法来解决速度子问题中因可变粘性和各向异性而带来的求解困难;对

欧洲大陆冰盖模型模拟的测试中,发现完全冰盖模型模拟的结果不存在简化模型的非物理解,并对格陵兰岛的冰盖进行了数值模拟,预测短期冰盖运动行为。



他们基于本所的 PHG 平台开发了冰盖完全模型的模拟程序,并在本所的机群上进行了测试,该程序在使用到上千个核并有近亿未知数的规模模拟时仍然具有很好的并行效率。目前,该程序已经被吸收为美国的天气气候模拟软件 MPAS 的一部分。跨尺度预测模式 (MPAS) 是研究天气,局部气候和全局气候模拟的项目,该项目的研究单位包括美国洛斯阿拉莫斯国家实验室的气候研究组和美国国家大气研究中心,该项目目前包括大气、海洋和陆地冰盖模拟三个组成部分,其中冰盖模拟的完全模型模拟部分使用了本所研究人员开发的三维冰盖求解器。

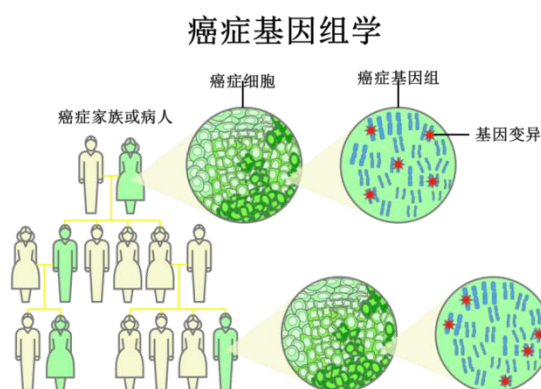
地球冰盖模拟是研究全球气候变化的重要一环,准确、高效的冰盖模拟是人们准确预测全球气候变化的关键之一,本研究发展冰层问题的完全模型的高效求解方法,使得冰盖完全模型的数值模拟得以实现。他们今后将继续冰盖模拟的

其他组成部分的研究, 如冰盖入海浮冰的断裂, 冰盖温度变化和冰盖底部含水层变化等等, 以求

更全面更准确的预测冰盖运动和融化。

计算癌症基因组研究取得进展

文: 生物医学部



高维数据的分析与建模对数学方法带来了极大的挑战。在生物医学交叉研究中, 高通量测序、芯片等实验手段的快速发展产生了海量的高维数据。特别是针对癌症这种复杂疾病, 国际上已经启动了若干大的项目, 构建了专门的数据库, 来产生和积累相关的数据。例如 2005 年 12 月由美国国家癌症和肿瘤研究所 (NCI) 和国家人类基因组研究所 (NHGRI) 联合启动的癌症和肿瘤基因组图谱 (The Cancer Genome Atlas, TCGA) 计划, 采用基因组分析技术, 特别是采用大规模的基因组测序, 将人类全部癌症 (50 种包括亚型在内的肿瘤) 的基因组变异图谱绘制出来, 并进行计算分析, 旨在了解癌细胞发生、发展的机制。该项目预计耗资 1 亿美元, 截止到 2012 年, 针对乳腺癌、脑癌等几个癌症类型, 积累了海量的高维基因组学数据, 怎样分析和利用这些数据为计算生物研究者提出了强有力的挑战。数学与交叉科学中心生物医学研究部的章祥荪、吴凌云、

王勇、张世华组成的研究团队针对具体问题, 发展了若干新的方法, 取得一系列成果。

(一) 高维基因组数据降维的最优化方法

癌症的发生、发展和预后受诸多遗传和环境因素的影响。正确选择癌症的分子标记物对研究癌症的生物学机理、最终指导系统综合治疗具有重要意义, 这也是癌症基因组学研究的一个主要目标。在数学上, 该问题抽象为高维基因组数据的降维问题, 即对这些数据进行高效的、保留较强生物解释性的降维, 这是计算癌症基因组研究的核心问题之一。值得一提的是, 构建将数据维数降低、利于进一步处理、可视化和理解信息的数学模型, 已成为各科学领域面对的一个共同问题。该问题在人工智能和机器学习领域也被称为特征选择, 是一个根据一定评估准则最优地从初始高维特征集合中选出低维特征集合的过程。

他们从最优化方法论的角度入手探求该问题的确定性模型和算法, 发展了基于线性规划的高维数据降维的方法论框架。具体地, 提出同时最大化数据解释能力和最小化特征数目的多目标优化模型, 进一步将该问题转化为单目标优化, 松弛优化变量用线性规划来近似整数规划, 为本质是 NP 难的组合优化问题设计了高效的求解算法。该方法已经成功地应用到基因组和代谢组学数据上。与中国医学科学院病原生物学研究所合作, 分析癌症病人的基因表达数据, 针对样本特

异性导致的强非线性, 引入椭圆分类器的概念, 能够选出非冗余的关键基因作为生物标记物, 取得了很高的分类精度。成果发表在生物化学和分子生物学类的知名期刊 *Nucleic Acids Research*。审稿人认为, “该新方法为高维数据的统计分析提供了一个快速的近似解法, 而传统的统计方法在维度增高时变得不可行”。此外, 基于该方法与成都中医药大学合作分析代谢数据研究针灸的分子机理, 成果发表在系统生物学顶级期刊 *BMC Systems Biology*。

XianwenRen*, Yong Wang*, Luonan Chen, Xiang-Sun Zhang and Qi Jin(*Co-first authors). ellipsoidFN: a tool for identifying a heterogeneous set of cancer biomarkers based on gene expressions, *Nucleic Acids Research*, 10.1093/nar/gks1288, 2012
Yong Wang, Qiao-Feng Wu, Chen Chen, Ling-Yun Wu, Xian-Zhong Yan, Shu-Guang Yu, Xiang-Sun Zhang, and Fan-Rong Liang. Revealing metabolite biomarkers for acupuncture treatment by linear programming based feature selection. *BMC Systems Biology*, 6(Suppl 1):S15, 2012.

(二) 多重基因组数据集成的模型和优化方法

随着多重组学数据的积累, 如何分析和利用这些数据, 特别是理解和认识这些不多类型特征之间的联系成为当前计算生物学最具有挑战性的课题之一。癌症组学数据的丰富积累, 为思考和研究该问题提供很好的可能和资料。

针对这一课题, 张世华博士与合作者提出了分析癌症多重水平组学数据的多重非负矩阵分解方法; 同时为了整合其他生物网络数据, 进一步提出了基于网络正则化的半监督学习的多矩

阵分解方法; 并考虑了稀疏性约束对模型的影响, 提出了高效的优化算法。他们还考虑了不同组学数据对基因表达的影响, 提出了考察整合多重数据对表达影响的偏最小二乘方法。他们还研究了驱动基因通路的快速识别算法。相关研究成果被发表在国际顶级杂志 *Bioinformatics* (3 篇, 包括 ISMB 一篇) 和 *Nucleic Acids Research*。

Shihua Zhang, Chun-Chi Liu, Wenyuan Li, HuiShen, Peter Laird, XianghongJasmine Zhou. Discovery of multi-dimensional modules by integrative analysis of cancer genomic data, *Nucleic Acids Research* 2012, 40(19): 9379-9391.

Wenyuan Li*, Shihua Zhang*, Chun-Chi Liu, Xianghong Jasmine Zhou.(*Co-first authors). Identifying multi-layer gene regulatory modules from multi-dimensional genomic data. *Bioinformatics* 2012, 28(19): 2458-2466.

Junfei Zhao*, Shihua Zhang*#, Ling-Yun Wu, Xiang-Sun Zhang.(*Co-first authors). Efficient methods for identifying mutated driver pathways in cancer, *Bioinformatics* 2012, 28(22): 2940-2947.

Shihua Zhang, Qingjiao Li, Juan Liu, Xianghong Jasmine Zhou. Integrating multiple functional genomic data to define microRNA-gene regulatory modules by a sparsenetwork-regularized multiple matrix factorization method. *Bioinformatics (ISMB2011)* 2011, 27:i401-i409.

(三) 癌症亚型发现的最优化方法

癌症并不是单一病症。以乳癌为例, 根据受激素影响情况的不同, 分为雌激素受体阳性和雌激素受体阴性两种类型。而根据所患肿瘤类型的不同, 乳癌患者接受药物治疗的效果也有所不同。通过计算方法分析癌症病人的数据, 发现、确定癌症的亚型, 并进一步找到癌症亚型在分子层面的标记物, 可以发展针对特定分子亚型的癌症疗法, 最终能帮助医生调整治疗病人的个性化治疗, 从而更有效地救治更多的癌症病人, 这是未来医学的发展方向之一。

从方法论来看, 该问题抽象为利用高维生物数据进行分类和聚类研究。在机器学习的方法论研究及针对生物数据的应用中, 分类和聚类一般是作为两个独立的问题来研究的。他们探究了

这两类的问题的内在联系, 提出了一个新的最优化模型, 既可以作为有监督的分类, 也可以进行无监督的聚类, 同时针对问题的结构, 他们构造了一个快速的迭代算法, 使得可以求解上万维度的问题。该方法应用在癌症基因芯片数据分析中, 取得了很好的效果, 成果发表在生物信息学顶级期刊 *BMC Bioinformatics*. 文章接收后, 被主编邀请精选到 *Cancer Bioinformatics* 系列中。而且论文发表后不久就成为该杂志下载次数最高 (Highly accessed) 的科学论文之一。

XianwenRen, Yong Wang, Jiguang Wang and Xiang-Sun Zhang. A unified computational model for revealing and predicting subtle subtypes of cancers. *BMC Bioinformatics* 13:70, 2012.

L^2 延拓问题和 Suita 猜想的解决

文: 物理/工程部

周向宇与他的学生关合作解决了 L^2 延拓定理中的最优常数问题, 作为推论解决了四十年前提出的著名的 Suita 猜想。

在此之前, 周向宇与他的两个学生关启安、朱朗峰合作提出了两个方法来处理最优常数问题, 并用其中一种方法得到比以往常数更优的常数估计, 这方面的前期工作在 *Comptes Rendus* 2011 及 *J. Math. Pures Appl.* 2012 上发表。

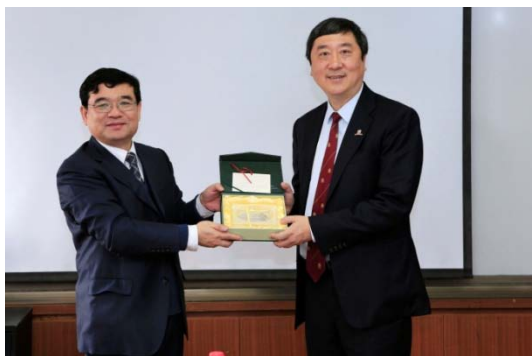
在综合比较、理解 Siu、Demailly、Berndtsson、Ohsawa 等人方法基础上, 提出了自己的方法; 发现了 Stein 流形上对 $\bar{\partial}$ 闭的流动形求解 $\bar{\partial}$ 方程的 L^2 估计的方法, 利用 current (流动形) 技巧直接给出 $\bar{\partial}$ 方程解在非光滑权情形下的 L^2 估计

(以往的方法都是考虑用光滑权来逼近非光滑权); 得到有一般非光滑扭曲元的 Bochner-Kodaira 公式 (以往只有光滑扭曲元或特殊非光滑扭曲元的 Bochner-Kodaira 公式), 并用于解 $\bar{\partial}$ 方程的 L^2 估计; 首次把最优常数问题最终转化成常微分方程求解问题 (以往的方法是先选取好函数再估计, 而周及其学生是设待定函数进行估计, 自然形成对待定函数的常微分方程, 再求解常微分方程以确定待定函数); 主要结果还减弱了 Ohsawa 对于可忽略权的限制条件, 即把 φ 和 ψ 的多次调和性减弱为 $\varphi + \psi$ 和 ψ 的多次调和性。

综合新闻

香港中文大学沈祖尧校长一行访问交叉中心

文：交叉中心办公室 图：王林



1月28日上午，交叉中心主任郭雷院士，数学院执行院长、交叉中心副主任王跃飞研究员等亲切会见了来访的香港中文大学校长沈祖尧院士一行。该校副校长程伯中教授、徐扬生院士、系统工程与工程管理学系主任蒙美玲教授、机械与自动化工程系主任黄捷教授、计算机科学与工程学系吕自成教授等参加了会见。

郭雷院士首先对沈校长率团访问交叉中心表示衷心欢迎。接着王跃飞执行院长简要介绍了中科院在科研体制、人才培养、国际交流等方面的情况及实现“三位一体”的未来发展定位；高小山副院长介绍了数学院和交叉中心的总体情况，并着重介绍了交叉中心的科研布局情况。

随后沈校长对中大的整体情况进行了介绍。他表示，不同领域对数学的需求很大，尤其是海量数据在不同领域的研究中发挥重要作用。通过此次访问，希望对交叉中心有更深刻系统地了解，挖掘梳理出关键问题，寻求双方的合作可能。

在友好热烈的气氛中，双方就在生物医学、传染病控制、经济预测、城市监控等大数据时代的挑战问题开展合作研究进入了探讨，并初步制定了行动计划。

郭雷指出，交叉中心将积极思考今后与中大开展合作研究的机制体制，并大力支持科研人员各种形式的交流互访。交叉中心具备良好的综合交叉实力和基础，覆盖的研究领域广泛，可以以6个研究部为依托，通过寻找共性问题，开展合作研究。

会谈结束后，双方互赠纪念品并合影留念。之前，沈校长一行还饶有兴致地参观了中科院科学与工程计算国家重点实验室的三号集群LSSC-III及中科院预测中心经济监测预警与政策模拟仿真平台演示。

参加座谈的还有崔俊芝院士，交叉中心副主任高小山研究员，交叉中心副主任闫桂英研究员，



系统所所长张纪峰研究员,应用所所长曹道民研究员,夏建明研究员,杨翠红研究员等。

背景链接:

香港中文大学成立于1963年,是香港成立的第二所大学,亦是香港8所受香港政府大学教

育资助委员会资助并可颁授学位的高等教育院校之一,由新亚书院、崇基学院、联合书院以及后逸夫书院组成,是香港唯一实行书院联邦制的大学。现设有八个学院:文学、工商管理、教育、工程、医学、理学、社会科学、法律学院。

“问题驱动的应用数学研究”研讨会召开

文:交叉中心办公室 图:王林/曹箭



2012年12月20日,由国家自然科学基金委(以下简称“基金委”)数理学部主办,中国科学院国家数学与交叉科学中心(以下简称“交叉中心”)、中国科学院数学与系统科学研究院(以下简称“数学院”)承办的“问题驱动的应用数学研究”研讨会在北京召开。来自国家基金委、中科院数学院、中科院相关研究所、北京应用物理与计算数学研究所、中国科学院大学、北京大学、南开大学、西安交通大学、山东大学、大连理工大学、北京交通大学、香港中文大学等多所高校、科研机构的多位院士、知名学者以及来自微软、宝钢等企业界的交叉领域专家70余人参加了此次研讨会。

交叉中心主任郭雷院士主持了开幕式。他首先对远道而来的各位专家致以诚挚的问候,他指

出,此次会议是一次创新尝试,旨在为数学家和非数学领域的专家提供一个交流合作平台,邀请非数学领域的优秀科技专家介绍重要实际问题对数学的需求,共同探讨具有重大实际背景的重要数学问题,以期共同推动问题驱动的应用数学与交叉科学的研究与发展,为重要实际问题的解决提供必要的数学理论和有效方法,并进一步促进不同领域的实质性交流与交叉。



数学院执行院长、交叉中心副主任王跃飞研究员代表数学院对各领域专家的出席表示感谢并预祝会议圆满成功。王院长在致辞中指出,此次会议虽然时间短暂,但可以以此为契机,在各领域专家之间建立起桥梁与纽带,探讨大家共同关心的科学问题和研究目标,为今后的进一步合作打下基础。



基金委数理学部雷天刚处长代表基金委感谢交叉中心和数学院承办此次会议。他指出,基金委非常支持问题驱动应用数学研究项目,该项目旨在进一步推动数学工作者与实际部门的合作,开展与其他领域结合的应用数学研讨,提升数学工作者面向国家重大需求从事应用数学研究的能力。

研讨会共安排了 11 个报告, 报告内容聚焦多个交叉领域中的数学问题。报告由中科院数学院执行院长、中科院交叉中心副主任王跃飞研究员、数学院副院长、交叉中心副主任高小山研究员、数学院计算数学所所长、交叉中心副主任陈志明研究员主持。中科院院士、大连理工大学程耿东教授, 中国科学院地理科学与资源研究所副所长周成虎研究员, 中国科学院动物研究所张知彬研究员, 中国科学院自动化研究所田捷研究员, 中科院微生物研究所张立新研究员, 西安交通大学齐春教授, 中国科学院院士、中国科学院生物物理研究所陈润生研究员、香港中文大学生物医学工程联合研究中心主任张元亨教授、中国科学院软件研究所丁丽萍研究员、宝钢中央研究院首席研究员郭朝晖教授、微软中国有限公司工程师任伟先后就结构拓扑优化及其推广、地理科学研究中的若干数学问题、生态网络建模及稳定性分析、医学成像中的数学问题、数学模型指导下的互动药物筛选、头盔瞄准中的图像与数学、与基因组研究相关的大规模计算、生物医学工程与健



康信息学领域核心问题对数学的重要需求、高安全等级操作系统研发中的数学问题、热轧带钢力学性能预报中的数据建模问题及互联网广告投放中的实际数学问题等做了精彩报告, 得到与会者普遍好评。

报告结束后, 与会专家就报告内容、体会与感受、交叉学科的发展、以及如何进一步推动问题驱动的应用数学的研究, 进行了广泛热烈地讨论, 并提出诸多设想与建议。

郭柏灵院士、彭实戈院士、鄂维南院士、陈润生院士等纷纷对邀请实际应用领域专家参与应用数学研讨会的形式表示肯定, 希望今后有更多机会深入了解这些领域对数学的需求。

从 2013 年起, 交叉中心将以研究部为单位, 定期召开交叉学科研讨会, 以期不断促进问题驱动的应用数学发展。



孙笑涛研究员、吕金虎研究员 分别获得 2012 年度国家自然科学二等奖

文：周宏

2013 年 1 月 18 日上午,在北京举行的国家科学技术奖励大会上,孙笑涛研究员完成的成果“模空间退化和向量丛的稳定性”及吕金虎研究员等完成的成果“若干新型非线性电路与系统的基础理论及其应用”分别获得 2012 年度国家自然科学二等奖。

孙笑涛研究员主要研究模空间退化、向量丛稳定性及其应用,取得如下成果:1)证明任意秩的广义 θ 函数空间的分解定理,被国际同行列为“退化技巧两个成功的应用之一”,称为“Narasimhan-Ramadas-Sun Approach”。论文被“美国数学评论”称之为“remarkable paper”;2)证明了 $SL(r)$ -丛模空间退化的 Seshadri-Nagaraj 猜想;3)发现并证明了 Frobenius 同态与稳定向量丛之间的重要联系,所得结果被国际同行称为:“一个重要公式”,“著名的不等式”,“众所周知的不等式”;4)证明了模空间中过一般点的极小有理曲线与 Hecke 曲线的等价性,国际同行在综述文章中用 3 个定理介绍了该结果和应用,所得到的引理成为其它研究的起点;5)与他人合作,证明“Arakelov-Yau 不等式等号成立”的 $K3$ 曲面纤维化的“modularity”,被列为建立“模猜想”(modularity conjectures)的 9 种方法之一。

吕金虎研究员与香港城市大学陈关荣讲座教授、广东工业大学禹思敏教授合作,主要研究

若干新型非线性电路与系统的理论设计、电路实现及其应用问题,主要贡献如下:1)证明了一类基本三维时滞多卷波 ($n>2$) 系统的混沌吸引子存在性,且该方法有普适性,可以类似拓广到其它高维系统,解决了非线性多卷波 20 多年来困扰理论界的基础性难题,被 J. Vandewalle 院士称为“里程碑”,并入选 2005 年 IEEE 电路与系统学会“电路与系统进展”;2)突破传统模拟电路的设计瓶颈,首次用模拟电路物理实现了单方向 14 卷波、双方向 $14*10$ 卷波和三方向 $10*10*10$ 卷波吸引子,创造了卷波数的模拟电路实验记录,解决了 20 多年来长期悬而未决的多方向 ($d>1$)、大数量 ($n\geq 10$) 多卷波物理实现的理论和技术难题;3)提出了三维空间中单方向、双方向和三方向饱和函数序列多卷波的设计方法并给出了理论证明和模拟电路实现,入选 2005 年 IEEE 电路与系统学会“电路与系统进展”;4)提出了广义 Jerk 电路、三维切换流形、多折叠环面和四维网格状多环面等一系列设计方法,相关技术获 6 项发明专利授权,其中关于广义 Jerk 电路的工作被 IEEE 电路与系统最高奖得主 L. O. Chua 院士称为给工程应用“建立了理论基础”;5)通过引入额外爆破点,提出了一种新的系统化电路设计方法并首次用模拟电路物理实现了单方向 10 卷波吸引子,被 J. Vandewalle 院士称为“里程碑”。

材料环境部联合学术研讨会第二次会议召开

文:刘颖



2012 年 11 月 4 日, 由材料环境研究部主办的材料环境部联合学术研讨会第二次会议在北京召开。材料环境部崔俊芝院士等 30 余位成员参加了研讨会。

此次研讨会主要议题为数学与能源领域的交叉研究。材料环境部能源专题陈志明、张林波研究员等 6 位成员做了精彩报告。他们分别就能源与计算领域各自研究工作的最新进展作了介绍, 并就未来可能合作的问题做了深入讨论。会上, 材料环境部成员达成共识, 在今后的研究中, 发挥科研创新精神, 探寻新的科研突破点, 更好的将基础性研究与前沿性、应用性研究相结合。

会议初步确定了材料环境部联合学术研讨会第三次会议的召开时间, 议题为计算数学与环境领域的交叉研究。

PDE 分析与计算研讨会召开

文:刘颖



2012 年 11 月 3 日, 由材料环境研究部与物理/工程研究部合作, 材料环境部主办的“PDE 分析与计算研讨会”在北京召开。物理/工程研究部曹道民研究员、材料环境研究部陈志明研究

员等 30 余位科研人员参加了此次会议。

物理/工程研究部张平研究员、材料环境研究部周爱辉研究员等 6 位科研人员分别就各自的交叉研究工作情况作了精彩报告。报告内容丰富, 涉及流体力学、磁流体力学、电子结构模型等诸多领域。会上, 两个交叉研究部的成员就将来可能合作的课题做了初步的探讨。

此次研讨会加强了材料环境研究部与物理/工程研究部两个交叉部的合作与交流, 为与会科研人员今后数学与其他学科交叉合作的高水平研究提供了新的思路与方法。

2012年计算机辅助制造、工程与数控中的数学与算法国际会议召开



2012年计算机辅助制造、工程与数控中的数学与算法国际会议于2012年10月25-26日在数学与系统科学研究院召开。会议邀请了美国、英国、意大利、法国、比利时、以色列等国家的大学与科研单位的,以及国内的中国科学院沈阳计算技术研究所、中国科学院沈阳自动化研究所、华中科技大学、大连理工大学、西北工业大学、清华大学等单位的50余位老师和研究生参加。

会议由先进制造部主任、中国科学院数学机

械化重点实验室主任李洪波研究员主持并致开幕词。以色列理工学院 (Technion Israel Institute of Technology) 的 Moshe Shpitalni 教授, 法国国家信息与自动化研究院 (INRIA Sophia Antipolis) 的 Bernard Mourrain 教授, 中国科学院数学与系统科学研究院副院长、国家数学与交叉科学中心副主任高小山研究员, 大连理工大学的孙玉文教授, 伊利诺斯州厄巴纳-香槟大学 (University of Illinois at Urbana-Champaign) 的 Placid M. Ferreira 教授, 西北工业大学的张卫红教授, 意大利帕多瓦大学 (University of Padova) 的 Alberto Trevisani 副教授等分别做了各自研究领域的大会报告。

此次会议促进了国内外数字化制造领域国内外科研单位之间的交流与合作, 对数学与先进制造领域学术研究的交叉与融合起到了重要作用。

“中澳量子控制会议”召开

文: 骆顺龙 图: 王林

由中国科学院数学与系统科学研究院、中科院国家数学与交叉科学中心与澳大利亚国立大学等联合发起和组织的“中澳量子控制会议”于2012年11月5日--8日在中国科学院数学与系统科学研究院顺利召开。这次会议得到国家数学与交叉科学中心的大力支持。国内外120余专家学者和研究生参加了该会, 外院代表90余位, 其中境外或国外代表17位。中国科学院郭雷院士

(会议主席), 澳大利亚科学院 Howard M. Wiseman 院士, Ian R. Petersen 院士, 北京计算科学中心主任林海清教授等参加了会议。

开幕式由组委会主席骆顺龙研究员主持, 数学与系统科学研究院副院长、国家数学与交叉科学中心副主任高小山致欢迎辞。

此次会议的主题是量子控制。这是一个量子物理, 控制论与信息论的交叉领域。会议安排了

34 个邀请报告和 13 个张贴海报 (poster)。会议具体议题包括量子反馈控制, 量子相干控制, 量子控制的极限能力, 量子测量与控制, 量子信息处理等。与会者就量子控制及相关领域的若干最新进展及国际前沿课题进行了活跃而又高层次的学术交流。

这次会议的召开, 为今后数学与系统科学研究院, 国家数学与交叉科学中心在量子控制领域与国内外其他机构开展广泛、深入的科研合作以及人才培养奠定了重要基础, 特别地, 进一步加强了与澳大利亚若干大学的合作。

中澳在控制领域有着长期的富有成果的合作关系。近年来, 在量子控制领域的合作关系尤为密切。特别地, 经中国科学院数学与系统科学研究院与澳大利亚国立大学等单位协商, 计划交



替在澳大利亚和中国举办系列量子控制会议, 其中第一届会议: Australia-China Quantum Control Workshop 2010 (澳中量子控制会议 2010) 已于 2010 年 11 月 1-3 日在澳大利亚国立大学召开, 并取得良好效果。此次北京中澳量子控制会议是该中澳合作计划的继续, 是该系列会议的第二届。

第二届国际油藏数值模拟研讨会召开

文: 张晨松



2012 年 12 月 6 日—7 日, 由中科院国家数学与交叉科学中心与中国石油勘探与开发研究院软件中心联合举办的第二届国际油藏数值模拟研讨会在翠宫饭店召开。来自大型石油公司 (中石油、中海油、美国 Chevron 公司、沙特

Saudi Aramco 公司)、油藏服务公司 (斯伦贝谢、默凯斯能源技术公司)、研究机构 (中科院计算数学所、软件所、过程所、地球物理所) 和高等院校 (北京大学、清华大学、中国石油大学、中国地质大学、北京科技大学、美国宾州州立大学、加拿大卡尔加里大学) 等单位的专家学者 80 余人参加了为期一天半的会议。

研讨会主题为新一代油藏数值模拟技术和挑战, 并分三个专题“大规模并行油藏模拟的模型、离散和求解技术”、“数值模拟在非常规油气田、低渗油田、页岩气和致密气田提高采收率技术中的应用”、“研制新一代油藏数值模拟器所面临的机遇和挑战”分别探讨。其中, 来自 Saudi Aramco 的 Ali Dogru 教授介绍了他们开发的、

目前世界最大并行规模的油藏模拟软件 GigaPOWERS, 利用 6000 核的并行机他们模拟沙特最大油田 10 亿多网格数的油藏; 这样的精细计算的结果为实际生产带来了巨大的效益。

与会代表对非常规油气田的渗流特点、油藏数值模拟器开发中存在的困难、新一代油藏数值模拟的设计与开发、油藏商业数模软件的现状、新型多尺度离散方法和代数方程组求解算法等问题展开了热烈地讨论。通过与地质工作者、油藏工程师的研讨, 与会人员对我国当前油藏工程应用单位的实际需求有了进一步了解, 并对数学



理论、数值算法及并行计算软件等方面的最新进展及其在油藏数值模拟中的应用前景有了更深刻地认识。

面向大规模科学计算的网格生成和可视化技术研讨会召开

文: 崔涛

2012年12月27日至12月28日, 由中科院国家数学与交叉科学中心材料环境研究部、科学与工程计算国家重点实验室和中国科学院数学与系统科学研究院计算数学与科学与工程计算研究所资助的“面向大规模科学计算的网格生成和可视化技术研讨会”在京召开。来自德国、西班牙以及国内高校、科研院所的60余名专家学者和研究生出席了研讨会。

研讨会的主题是大规模网格生成和并行可视化技术研究及软件研制。近年来, 随着高性能计算机的飞速发展, 进行数十上百亿未知量的超大规模数值模拟已不再是梦想。大规模数值计算除了给数值算法带来巨大挑战, 也对计算的前后处理系统提出了更高的要求, 其中尤以网格生成和计算结果可视化最为。网格生成和计算结果

并行可视化技术已成为大规模数值计算的瓶颈之一。研讨会上, Tetgen 软件包研制者和德国维尔斯特拉斯应用分析和随机研究所研究员 Si Hang 博士, 西班牙工程数值计算方法国际中心 GID/Kratos 软件研发团队负责人 Nelson Lafontaine 博士、流程工业综合自动化国家重点实验室副主任王成恩教授、计算数学与科学与工程计算所卢本卓研究员等 10 位学者分别就各自研究方向的最新进展作了介绍, 并就未来可能发展方向进行了展望。

此次研讨会旨在将网格生成和并行可视化技术领域的专家、算法研究专家以及应用领域专家联系在一起, 搭建交流和沟通的平台, 推动交叉研究。与会专家表示, 此次研讨会非常有意义, 希望该研讨会能办成一个系列会议。

学术动态

交叉中心报告论坛之二十四： 俄罗斯科学院院士 Nikolay S. Zefirov 谈化学结构与数学

图/文：交叉中心办公室



2012 年 12 月 7 日下午俄罗斯科学院院士、国际数学化学科学院院士、莫斯科国立大学杰出教授 Nikolay S. Zefirov 教授应交叉中心的邀请到国家数学与交叉科学中心访问并做了综合报告，报告题目为 Computer search for novel

reactions using "Formal-Logical Approach"。生物医学部副主任李雷研究员主持报告会并为 Nikolay S. Zefirov 院士颁发了讲座证书。

Nikolay S. Zefirov 院士深入浅出地介绍了一些化学结构之间相互联系的发现方法，数学学科中图论与一些化学拓扑结构的关系，以及未来可能交叉研究的重点，最后细致热情地回答了听众的提问。

Nikolay S. Zefirov 院士出生于 1935 年，1966 年获得前苏联莫斯科国立大学博士学位。他是俄罗斯科学院院士、国际数学化学科学院院士、莫斯科国立大学杰出教授，他的研究领域主要涉及有机化学、数学化学和计算机、医学化学等。

交叉中心报告论坛之二十三： 浙江师范大学朱绪鼎教授谈 “Graph colouring model for scheduling problems”

文：蔡晓宇 图：曹箭

2012 年 11 月 29 日上午，国家“千人计划”入选者、浙江师范大学教授朱绪鼎受交叉中心信息技术研究部邀请，在思源楼做了题为“Graph colouring model for scheduling problems”的报告。这是国家数学与交叉科学中心第二十三场综合

报告会。来自中科院数学院、中国科学院大学与清华大学的 30 余名师生聆听了报告。

该场报告旨在建立图的染色模型与排序问题之间的联系，以期图的有关理论可以应用于实际应用领域。朱绪鼎教授首先给出了排序问题的

图染色模型,介绍了相关的图的理论基础和近年来的一些综合性成果,最后提出了一些可以研究的问题和猜想,以及可能进一步研究的新的结合点。

朱绪鼎教授 1959 年 4 月出生于湖北仙桃,曾担任台湾中山大学西湾讲座教授,第三批国家“千人计划”入选者。作为图论与组合优化研究领域的著名专家,朱绪鼎在图的染色理论、结构分析、演算法等领域做出了杰出的贡献,取得一系列重大研究成果,特别是创新发展了图的圆染色理论,使得该方向已成为当前国内外研究的热点之一。朱绪鼎曾先后在国际重要学术刊物上发表



SCI 研究论文 130 余篇,论文被同行引用 1000 余次。2007 年,在 ISI 公布的、有全球 900 多位数学家入选的世界数学家被引用次数排名中,位列第 67 名。

交叉中心报告论坛之二十二： 美国斯坦福大学刘太平教授谈动力学理论研究

文：黄飞敏 图：曹箭



2012 年 11 月 14 日下午,交叉中心综合报告会第二十二场在数学院思源楼举行。台湾中央研究院院士、前中央研究院数学研究所所长、美国斯坦福大学教授刘太平受国家数学与交叉科学中心邀请,作题为“Kinetic Theory and Gas Dynamics”的综述性报告。国家数学与交叉科学

中心 100 余位科研人员及学生出席报告会。报告会由中心副主任高小山研究员主持。

在描述气体的运动时,当考虑不同的物理尺度或关注运动的不同侧面时可以得到许多著名的运动方程。这些方程大部分是经典的,可以追溯到十九世纪或以前。在宏观层次,最著名的方程是气体动力学方程(Gas Dynamics)中的 Euler 方程。在微观层次,相应的运动模型由描述单个粒子运动耦合的 Newton 方程构成。介于两者之间的一个介观模型是动力学理论(Kinetic Theory)中的基本方程,即 Boltzmann 方程。刘太平教授首先回顾了上述两个研究领域的历史、重要问题及取得的重要进展,如关于 Boltzmann 方程,评价了目前流行的 2 类研究方法,即统计力学方法

和流体力学方法；关于 Euler 方程，介绍了著名的 Prandtl 猜测及关于此猜测产生的大讨论。接着介绍了两个领域之间的紧密联系，如 Boltzmann 方程到 Euler 方程的流体动力学极限属于 Hilbert 第 6 问题的一部分等。最后列举了一些重要的公开问题。

刘太平现任中央研究院数学研究所特聘研究员，中央研究院院士，2000-2012 年任中研院数学研究所所长。主要研究方向为非线性偏微分方程、激波理论及动力学理论等。在 2002 年的国际数学家大会上做 45 分钟邀请报告，是偏微分方程领域的权威专家，在国际上有重要影响。

交叉中心报告论坛之二十一： 美国 Texas A&M 大学陈巩教授谈风能研究

文：姚鹏飞 图：曹箭

2012 年 11 月 6 日上午，交叉中心第二十一场综合报告会在数学院举行。美国 Texas A&M 大学的陈巩教授受国家数学与交叉科学中心邀请，作题为“Progress in the Mathematical Study of Wind Turbine Flows and Wind Energy”的报告。

风能是现代可持续能源的主要分支，陈巩教授在报告中详细介绍了他们在风能研究方面的工作，包括涡轮气流的控制，涡轮塔结构振动问题的建模及近海涡轮塔的结构等。他还给听众演示了由计算流体力学作出的涡轮气流的运动。

陈巩教授于 1950 年出生于台湾，1972 年毕业于台湾清华大学，1977 年获美国 Wisconsin 大学博士学位。他的研究兴趣包括分布参数系统控制，偏微分方程，偏微分方程数值解，工程力学，量子力学等。他是美国 Journal of Mathematical



Analysis and Applications 主编，并担任多个杂志的副主编，包括 the SIAM Journal on Control and Optimization, the International Journal on Quantum Information, 和 the Electronic Journal of Differential Equations。

数学文摘

诺贝尔奖为何青睐交叉学科

作者：冯一潇

在近万个独立学科中，一半左右属于交叉学科。目前比较成熟的学科大约有 5550 门，其中交叉学科总数约 2600 门，占全部学科总数的 46.8% 之多，其发展表现出良好势头和巨大潜力。

“大科学”并不是指依赖于大装置的“大科学”，而是指“综合性的大科学思维体系”，以区别于“传统的、狭隘的科学思维方式”。

冯一潇

百年诺贝尔奖，有 41.02% 的获奖者属于交叉学科。尤其在 20 世纪最后 25 年，95 项自然科学奖中，交叉学科领域有 45 项，占获奖总数的 47.4%。这个统计数据的重要意义，尤其值得不擅长多学科交叉的中国科学家深入思考。

交叉学科是指两门或两门以上学科融合而形成的一种新的综合理论或系统学问。交叉学科不是多门学科的简单拼凑堆积，而是多学科依存于内在逻辑关系联结渗透形成的新学科。不同的学科彼此交叉综合，有利于科学上的重大突破，培育新的生长点，乃至新学科的产生。

中国科学院院长路甬祥提出：在近 100 多年里，交叉科学，包括边缘科学、横断科学、综合科学和软科学等，运用多种学科的理论和方法，消除了各学科之间的脱节现象、填补了各门学科之间边缘地带的空白，将分散的学科综合起来，从而实现科学的整体化。交叉学科研究正在成为

科学发展的主流，不仅活跃研究者的思维，开阔科学研究的视野，同时也大大推动着科学技术的发展。一个最新的例子是纳米科学，这是最为典型的交叉学科，它的出现，推动了众多学科领域的发展。

学科交叉是重大科学成就的源泉

将某一学科已发展成熟的知识、方法和技术应用到另一学科的前沿，能够产生重大创新成果，学科交叉是创新思想的主要来源之一，已经取得了杰出的成就。以诺贝尔奖为例，交叉学科所获奖项一直占据很大比重，评奖委员会更倾向于表彰属于交叉学科范畴的研究成果。

20 世纪自然科学最重要的有三大发现，相对论、量子力学和 DNA 双螺旋结构。其中 DNA 双螺旋结构的发现是科学史上最富传奇性的“章节”之一，作出重大贡献的科学家一共有 4 位：物理学家克里克(Crick)和威尔金斯(Wilkins)、生物学家沃森(Watson)还有化学家富兰克林(Franklin)。他们 4 人具有不同的知识背景，在同一时间都致力于研究遗传基因的分子结构，在既合作又竞争，充满交流和争论的学术氛围中，发挥了各自专业的特长，为 DNA 双螺旋结构的发现作出了杰出贡献，这是科学史上由多门学科交叉渗透、相互借鉴产生的一项举世瞩目的科学成果，成为生命科学发展的重要里程碑。基因工

程中的DNA重组技术也是学科交叉的产物。这一技术的开拓者和创始人,美国生物化学家保罗·伯格(Paul Berg)借助类似工程设计的方法,利用限制酶和连接酶处理SV40病毒和大肠杆菌DNA碎片,最终两个不同来源的DNA片段连接在一起并发挥其应有的生物学功能。这是世界上首次完成的基因重组和DNA人工转移的重大创新研究,证明了完全可以在体外对基因进行操作,从而为人类主动改变生物的性状和功能,创造更加适合于人类需要的新生物提供了重要方法,开创了遗传工程的新纪元。

1998年度诺贝尔化学奖的颁布,向人们展示了数学、物理和化学学科的交叉和融合取得的重大成果。美国物理学家瓦尔特·科恩(Walter Kohn)和英国数学家约翰·波普(John Pople)以物理和数学工具,发展了量子化学理论和计算方法,在化学领域取得了骄人成就。通过以科恩和波普为代表的量子化学工作者的不断努力,今天,量子化学无疑成为化学工作者最有用的工具之一。磁共振成像技术(MRI)的发明实质上是物理学与医学的结合,也是交叉学科能产生丰富成果的有力证明。这种能精确观察人体内部器官而又不造成伤害的影像技术,对于医疗诊断、治疗及其检查至关重要。其发明者,美国的保罗·C·劳特伯(Paul·C·Lauterbur)和英国的皮特·曼斯菲尔德(Peter Mansfield)因此项技术获得了2003年诺贝尔生理学或医学奖。

交叉学科研究是科学发展的主要方向

交叉学科是当今科学的前沿研究领域。据统计,在近万个独立学科中,一半左右属于交叉学科。目前比较成熟的学科大约有5550门,其中交叉学科总数约2600门,占全部学科总数的

46.8%之多,其发展表现出良好势头和巨大潜力。当今,很多热门话题都涉及交叉学科研究,如基因组学与蛋白质组学、神经系统科学、微阵列技术等;同样,许多重大的科研成就也都是跨学科合作的成果,如人类基因组测序、“绿色革命”以及载人空间飞行等。前沿学科在交叉融合中获得新生。2003年由美国硅谷产学研各界组成的智囊机构一致认为,20世纪90年代迅速发展的生物技术、信息技术、纳米技术正在共同酝酿下一个科技创新高潮。未来这三大技术的交叉融合将有望广泛改变工艺和产品,产生新的经济增长点,形成一次新的产业革命,并将对全球产业产生重大影响。可以说,交叉学科研究是未来科学发展的主要方向,实现科学研究的跨学科性不仅是科学自身发展的需要,也是全球经济和人类社会发展的需要。

世界各国对交叉学科研究极为重视。2002年美国国立卫生研究院(NIH)建立了多个学科交叉研究中心,便于来自不同学科背景的科研人员相互交流和沟通,不仅设立了“多学科交叉研究人员培训基金”,还举办了“技术方法创新研讨会”和“生命科学与物质科学交界的机构联席会”。英国等其他发达国家也相继成立了学科交叉研究中心,为前沿学科建设开辟道路。1998年,诺贝尔物理学奖获得者,斯坦福大学教授朱棣文倡导确定了“生物学交叉学科研究计划”(Bio-X Program),包含了涉及生物科学、生物工程和医学领域的众多学科。Bio-X计划将基础理论、应用研究和临床科学的前沿结合在一起,促进整个生物医学领域从分子尺度跨越到人类器官尺度的技术创新。2002年,《自然》杂志设立了一项“交叉学科进步奖”提名,奖励那些在交叉学科领域作出杰出贡献的科学家,宗旨是为了促进学科

之间的交流。这吸引了大批来自不同学科科学家,他们有的在交叉学科领域中发明了新的技术,有的则将一些已经成熟的技术以新的形式去解决问题。大家都希望可以在思想上交融并闪现出智慧的火花。生命科学权威杂志《细胞》也特设了一个相关交叉学科领域的介绍专栏,为研究人员打开了进入交叉科学研究的大门。

我国政府在“十一五”科学技术发展规划中,选择了一批重大科学前沿问题开展研究并给予资助,包括:生命过程的定量研究与系统整合、凝聚态物质与新效应、地球系统过程与资源、环境和灾害效应等多学科研究方向。“973”项目8个资助领域之一的“综合交叉”,通过鼓励不同学科领域间的交叉融合,培养和造就一大批富有创新精神的科技人才,建设若干学科交叉、综合集成、机制创新的重点实验室和研究基地,从而孕育出更多的创新性研究成果。第一批“973”项目——“光合作用高效光能转化机理及其在农业中的应用”就是一项多学科交叉,把生物学、物理学、化学和农学有机结合开展科学研究的项目,使我国光合作用机理与膜蛋白三维结构研究处于国际领先水平。这样大跨度的多学科结合具有鲜明的特色在国际上也很少见。2006年,科学出版社出版的《中国交叉科学》创刊,为国内首个交叉科学研究连续出版物,以交叉为特色,反映我国交叉科学研究新思想、新观点、新成果。同年北京大学成立前沿交叉学科研究院,该院已建立了生物医学跨学科研究中心、化学基因组学研究中心等若干前沿交叉学科研究群体,承担国家重大科研项目。国内其他重点高校如浙江大学也相继成立了相关研究中心,加强交叉学科的建设。中国科学院高能物理所于2006年设立了有200多名研究人员的多学科研究中心,迄今已在

纳米生物安全性、金属蛋白质组学、分子影像学、生物大分子结构等方面取得了一批重要成果。

交叉学科研究的创新建设

不同学科间的交叉和融合是21世纪科学发展的主要趋势。交叉学科研究既有一般科学研究的共性又有自己的个性。突破传统思维方式和体制,促进各学科协作发展的关键是建立一套行之有效的组织管理机制、人员聘任评价机制、资源分配机制以及相应的交叉学科学术支撑体系。交叉学科研究应向体制化、制度化及管理的可操作化迈进。2004年,美国科学院协会(American Scientific Affiliation)发表了长篇报告,名为《促进交叉科学研究》(Facilitating Interdisciplinary Science)。该报告全面深入地分析了交叉学科研究的发展现状,对如何促进交叉学科发展提出了从科学研究、人才培养到管理体制一揽子富有创见性的改革方案,其前瞻性、系统性、深刻性令人深受启发。

与发达国家相比,中国交叉学科的发展还处在不成熟的阶段。交叉学科既没有院系对应,自身也未成学统,从事交叉学科的研究人员只能分散到其他学科。由于交叉学科尚没有专门的学术建制,因而常会遭受传统思维的束缚和固有研究体制与模式的阻力,找不到对应的位置,得不到政策、制度的鼓励和保障,致使学科交叉研究的力度和广度不足。《中国交叉科学》主编刘仲林教授认为,交叉科学研究要在现有科研和教育体制中立足,就必须从深层突破以传统学科界限为基础的科研管理和学科组织模式,建立有利于交叉、开放和共享的运行机制,拓展科研和教学充分自由创造的空间。正如李四光先生所言,要“打破科学割据的旧习,作一种彻底联合的努力”。

建设和发展交叉学科的关键取决于构建基于学科交叉的教学科研管理平台。应根据自身学科建设和发展情况, 积极创新平台建设模式, 以优势学科和特色学科为基础, 组织联合相关的研究力量, 逐步克服现行的教学科研体制、人事制度与学科交叉平台建设相互矛盾的突出问题, 建设一个学科前沿性与学科交叉性相结合、实体与虚体相结合的交叉学科研究平台。政府和资金赞助部门应有导向性, 以鼓励学科交叉研究的开展。通过优化资源的配置, 如设立交叉学科研究项目等, 保证资源合理、有效到位, 以较好地实现各要素间的优化配置, 并使各要素间得以和谐发展。还要构建学术交流平台, 不断创新学术交流方式, 通过不同学术观点的争鸣和学术思想的碰撞、切磋、互相渗透和融合, 扩大视野, 有效沟通, 启迪科技人员的学术思想, 激励创造性思维能力。

交叉科学则因其独特的跨学科性而无法找到自己的应有位置, 以至形不成自己的专门队伍和组织, 缺乏人才培养基地和学术研究平台, 发展不得长时间停留在“业余”的水平上。科技体制应为交叉学科发展培养创新人才, 整合人才资源, 组建科研团队。利用不同学科背景的研究者在方法上的相互启发, 科研中的相互配合, 真正实现质量上的系统整合效应。通过选拔优秀人才、激励先进成果来推动科技进步。逐步建立和完善交叉学科建设的弹性绩效考核和评价机制, 弱化短期考核指标, 努力为学科建设特别是交叉学科的建设和发展创造良好的软环境。正如美国国家科学技术委员会 2004 年给美国政府的科技咨询报告中指出: “为使未来科技领域真正有所成就, 造就能够跨越传统学科进行研究并思考外部世界的新一代科学家是绝对必要的。培育这种要么能够跨学科研究, 要么知道如何在学科交叉领域

与他人合作的新一代研究人员, 对于未来至关重要。”

学科交叉推动“大科学”时代进程

学科交叉体现了科学的综合化发展趋势。科学上新理论、新学科、新技术的出现常常是在现有学科的边缘或交叉点上。学科交叉已经形成了大量成熟的交叉学科, 如物理化学、生物化学、生物物理学等, 并且还将陆续形成其他重要的交叉学科。这些新学科将会大大推动科学进步, 尤其在现代“大科学”发展趋势下, 几乎已找不到没有学科交叉的纯科学问题。随着学科交叉研究的发展, 新兴交叉学科的产生以及各种新的理论体系和研究方法的创建与不断完善, 使得科学本身向着更深层次和更高水平发展, 推动科学向着多维综合性、创造性和开放性的思维方式迈进, 这就是所谓的“大科学”时代。我们这里所指的“大科学”并不是指依赖于大装置的“大科学”, 而是指“综合性的大科学思维体系”, 以区别于“传统的、狭隘的科学思维方式”。

“分化”是小科学时期科学发展的主要动力, “交叉”则是大科学时代科学发展的主要表征。在大科学背景下, 交叉学科得到突飞猛进的发展, 受到政府和学界的广泛关注。国务院 2006 年发布《国家中长期科学和技术发展规划纲要(2006—2020 年)》, 部署了四项重大科学研究规划, 涉及蛋白质研究、量子调控研究、纳米研究和发育与生殖研究等研究领域, 其中的 16 个重大专项多数都是跨学科领域的。中国科协新批准的 34 个学会中, 有一些就是交叉科学学会。可见, 交叉学科的不断涌现, 并逐渐形成各种交叉学科群, 乃是当代大科学时代的主要特征之一。

国家自然科学基金委员会于 2000 年开始试

点实施“重大研究计划”，这是提高我国科技持续创新能力的一项新举措，是重要的制度创新。该项目针对重大科学问题，整合不同学术思路 and 不同层次的项目形成具有统一目标的项目群，实施相对长期（6~8年）的支持，旨在促进学科交叉和学术争鸣，激励创新。这种资助模式强调顶层设计，突出战略性，在围绕整体目标进行研究方面作出了积极探索，既保证了科学研究自由探索的需要，发挥了科学家的创造性，又顺应了“大

科学”时代科学研究的规模性，实现了多学科的交叉和集成。基础研究中的创新常常来自不同学科之间的互碰、不同理论的互融和不同专业人员的互补。“大科学”时代将会越来越重视科学体系的统一性和完整性，这就是现今科学研究的显著特点和发展趋势。我国交叉学科目前总体上处在发展成长阶段，并且表现出快速发展的趋势，具有良好的发展前景和广阔的发展空间。

（来源于：科学网 发表时间：2012-12-06）

美国联邦政府启动“大数据研究与发展计划”

2012年3月29日，美国总统奥巴马宣布启动“大数据研究与开发计划”(Big Data Research and Development Initiative)，旨在提高从海量数字数据中提取知识和观点的能力，从而加快科学与工程发现的步伐，加强美国的安全，实现教育与学习的变革。

1. 计划推出的背景和意义

1) 背景

近年来，美国一直很重视大数据的研发。多个联邦机构启动了相关项目，涵盖国防、能源、航天、医疗等各个领域，但这些项目没有得到很好的协调。

2011年，美国网络与信息技术研发计划(NITRD)设立“大数据研发高级指导小组”，负责确定大数据国家计划的目标和跨联邦机构的大数据研发活动。此次大数据计划的推出正是该小组一年多研究工作的体现。

2) 意义

目前，企业界是发展大数据的主力军。此次

计划的推出表明，美国已经从国家战略高度来认识大数据并开始行动。白宫科技政策办公室的主管将大数据计划与历史上对超级计算和网络的投资相提并论，“过去在信息技术研发方面的联合投资推动了超级计算机和互联网的创建，而大数据计划有望利用大数据进行科学发现、环境和生物医学研究、教育以及保护国家安全的能力发生变革。”

2. 启动的大数据项目

为启动大数据研发计划，以美国国家科学基金会(NSF)为首的6大联邦机构宣布将投资超过2亿美元资助新项目的研发，其中国防部和国防部高级研究计划局(DARPA)都投入巨资，可见大数据对于国防安全的重要意义。此次启动的大数据项目包括：

1) 美国国家科学基金会与国立卫生研究院(NIH)

两机构联合招标的“促进大数据科学与工程的核心技术”项目将促进对大规模数据集进行管

理、分析、可视化并从中抽取有用信息的核心科学技术的发展。NIH 尤其关注与医疗和疾病有关的分子、化学、行为、临床等数据集。根据两机构发布的申请指南,申请提案应关注以下三方面:

①数据收集与管理

处理多源、异构、复杂的海量数据需要开发新的方法与工具。可能的研究领域包括但不限于:

——针对不断产生的数据以及共享和广泛分布的静态实时数据的新的数据存储、I/O 系统和架构;

——计算、存储和通信资源的有效使用与优化;

——能持续收集和处理数据,并确保其精确性、可信性和完整性的容错系统;

——能利用语义和情境信息自动注释数据的新方法;

——面向先进数据架构(包括云)的新设计,可以解决极限容量、电源管理、实时控制等问题,同时确保可扩展性和可用性;

——新一代多核处理器架构,以及能最大程度地发挥该架构优势的新一代软件库。

②数据分析

数据分析、仿真、建模和注释领域的进展将产生重大影响,有助于促进科学发现,认识事件的因果关系、进行预测并提出行动建议。可能的研究领域包括但不限于:

——新的算法、编程语言、数据结构和数据预测工具的开发;

——理解海量数据集计算的重要特性所需的计算模型和基础数学与统计学理论;

——针对不断产生的数据集的实时处理技术,以及允许更灵活、更直观地研究数据的实时可视化和分析工具;

——能整合不同数据并将数据转化为知识,

实现实时决策的技术。

③e-Science 合作环境

综合的“大数据”网络基础设施必不可少,它能使广大的科学家和工程师团队访问多样化的数据,以及最优秀、最实用的推理和可视化工具。可能的研究领域包括但不限于:

——有助于不同领域、不同地域的科研人员和学生互相协调工作,并大幅提高科学合作效率的新的合作环境;

——通过机器学习、数据挖掘和自动推理等方式实现科学发现过程的自动化;

——能管理多学科领域复杂和大规模科学成果流的新的数据管理技术;

——能促进科学工作流和新应用开发与使用的端对端系统。

除了关注以上三方面外,提案还必须包括能力建设方案,因为这对新兴科研教育领域的健康发展至关重要。此外,提案还可选择在某个优先领域开展大数据项目,例如医疗 IT、应急响应、清洁能源、网路学习、材料基因组、国家安全、先进制造等领域。

2) 国家科学基金会

国家科学基金会正在实施一项全面的长期战略,包括从数据中获取知识的新方法、管理数据的基础设施、教育和队伍建设的新途径,尤其是:

①鼓励科研院校开展跨学科的研究生课程,以培养下一代数据科学家和工程师;

②向加州大学伯克利分校提供 1,000 万美元的资助,将机器学习、云计算、众包这三种方法整合起来,用于将数据转变为信息;

③为“EarthCube”提供首轮资助,使地学家可以访问、分析和共享地球信息;

④向一个培训小组分配 200 万美元,使本科

生能在利用图形和可视化技术处理复杂数据方面获得培训;

⑤向一个由统计学家和生物学家组成的科研小组提供 140 万美元的资助,以确定蛋白质结构和生物学通路;

⑥召集跨学科的研究人员以确定大数据如何改变教学。

3) 国防部

国防部(DOD)每年投资 2.5 亿美元(6,000 万用于新的研究项目)资助开展以下研究:

①“数据到决策”,开发计算技术和软件工具,以分析那些与动态推理和推理机相连的海量数据(包括表格等半结构化数据和文本等非结构化数据);

②自动化,利用“数据到决策”取得的进展来开发相关的支持工具,这些工具能够识别趋势、适应现实世界的条件,并可不依赖人为干预而在复杂的动态环境中成功运行;

③人机系统,促进人机接口的发展,以实现运行和培训方面的无缝合作。

此外,美国国防部高级研究计划局开始启动“XDATA 项目”,拟在未来四年每年投资 2,500 万美元,开发分析大规模数据的计算技术和软件工具。项目拟解决的中心挑战包括:

——开发可升级的算法,以处理分布式数据仓库中的不完全的数据;

——创建有效的人机互动工具,使其可以根据不同的任务进行轻松定制;

XDATA 项目将支持开源软件工具包,为用户提供可在多种环境中进行大规模数据处理的灵活的软件。

4) 国立卫生研究院

NIH 的千人基因组计划数据集将通过亚马逊网络服务免费对外开放。这些数据总量达到

200 TB,是世界上最大的人类基因变异数据集。

5) 能源部

能源部(DOE)将提供 2,500 万美元的资助,建立“可扩展的数据管理、分析和可视化研究所”(SDAV)。SDAV 将汇集美国 6 个国家实验室和 7 所大学的专家,开发新的工具来帮助科学家管理和可视化来自 DOE 超级计算机的数据。

6) 地质调查局

地质调查局(USGS)的约翰·韦斯利·鲍威尔分析与集成中心启动了 8 个新的研究项目,以将地球科学理论的大数据集转变为科学发现。

3. 目前已经或正在开展的大数据项目

目前美国多个联邦政府部门正在开展多个大数据相关项目,以下重点介绍其中的部分项目。

1) 美国国防部高级研究计划局

①“多尺度异常监测”(ADAMS)项目:旨在解决海量网络数据中的异常监测和鉴定问题。

②“内部人网络威胁”(CINDER)项目:旨在开发创新的方法,以监测军事计算机网络中的网络间谍活动,并提高监测的精确度和速度。

③Insight 项目:旨在开发一个资源管理系统,通过分析来自成像和非成像传感器及其它来源的信息,自动确定网络威胁和不规则战争,从而弥补当前情报、监视与侦察系统存在的不足。

④“机器读取”项目:旨在开发能够处理自然语言的学习系统,并将由此产生的语义表征插入到知识库中,实现人工智能应用,而无须依赖于专家。

⑤“心灵之眼”项目:旨在为无人系统研发“可视智能”的能力,使无人系统通过对系统采集到的流媒体数据进行分析,发现并及时报告重要的作战信息,从而使作战人员能够及时地采取相应措施应对发生的重要事态。

⑥“面向任务的弹性云”项目：旨在开发监测、诊断和应对网络攻击的技术，解决云计算的安全挑战。

⑦“加密数据的编程计算”(PROCEED)项目：旨在针对那些在使用过程中保持加密状态的数据，开发实用的计算方法和编程语言，从而克服云计算环境中的信息安全挑战。由于无需在用户端解密数据，因而网络间谍也难以得逞。

⑧“视频与图像检索与分析工具”(VIRAT)项目：旨在为军事图像分析家开发一个系统，使其能利用收集到的海量视频内容，在某些事件发生时即可发出警报。VIRAT 还将开发相关工具，使分析家能够从超大型视频库中快速检索视频内容。

2) 国土安全部(DHS)

“可视化和数据分析卓越中心”(CVADA)通过对大规模异构数据的研究，使应急救援人员能够解决人为或自然灾害、恐怖主义事件、网络威胁等方面的问题。

3) 能源部

①科学办公室

——“先进科学计算研究办公室”(ASCR)带领数据管理、可视化和数据分析团体开展相关研究，包括 Kepler 等广泛使用的数据库管理技术、存储资源管理标准、ADIOS 等数据存储管理技术、ParaView 等科学可视化工具。

——高性能存储系统(HPSS)：DOE 和 IBM 等联合开发管理 P 级数据的软件，被纳米、核能物理、气候科学等学科及数字图书馆广泛利用。

——“千万亿次级数据的分析数学”项目：为从海量科学数据集中提取知识，发现重要特征和理解这些特征间的联系，该项目旨在解决与之相关的数学挑战，包括流数据的实时分析、随机非线性数据简化技术，以用于来自于电网的传感

器数据、气候数据等。

——“下一代联网”项目：支持开发有关发现、移动和利用大规模数据的科研合作工具，相关产品包括 Globus、GridFTP、ESG。

②基础能源科学办公室(BES)

BES 的“科学用户设施”项目支持帮助用户管理和分析大数据，如“加速数据获取、简化和分析”(ADARA)项目旨在解决散裂中子源数据系统的数据流需求，以为实验控制提供实时分析能力。

③聚变能源科学办公室(FES)

“先进计算带来科学发现”(SciDAC)旨在解决与聚变能源科学计算研究和实验研究相关的大数据挑战，其开发的数据管理技术和可视化技术受到欧盟和国际热核聚变实验反应堆(ITER)的关注。

④高能物理办公室(HEP)

“计算高能物理”项目支持有关大规模、复杂实验数据集以及模拟数据的分析研究。合作的大数据管理活动包括 PanDA 工作量管理系统和 XRootD(可快速访问多种数据集的高性能、容错软件)。

⑤核物理办公室(NP)

“美国核数据项目”(USNDP)旨在编写和交叉检验与原子核重要特点相关的所有实验结果。

⑥科技信息办公室(OSTI)

OSTI 是美国唯一参加国际 DataCite 的联邦机构成员，在制定数据引用的政策和技术方面发挥着重要的作用。它帮助实现数据的有效再利用与验证，并跟踪数据的影响。

4) 食品与药品管理局(FDA)

①“虚拟实验室环境”将整合现有的资源和能力，建立一个虚拟的实验室数据网络。

②美国国家档案记录管理局的“十亿电子记

录网络基础设施”(CI-BER)是由多个机构联合资助的测试床,其将评价相关的技术和方法,以支持对超大规模数据集的持续访问。

5) 国家宇航局 (NASA)

①为支持未来的地球观测任务, NASA 的信息系统必须不断演变。“先进信息系统技术”(AIST)项目旨在降低 NASA 信息系统不断演变的风险和成本,提高科学数据的可用性和利用率。

②“行星数据系统”(PDS)是有关 NASA 执行行星任务过程中所产生的数据产品库,其中的所有产品都经过同行评估,并可通过在线目录系统轻松进行访问。

6) 国立卫生研究院

①“癌症成像档案”(TCIA)是一项图像数据共享服务,旨在推动医学成像领域的科学发展。

②“人脑连接”项目旨在绘制人脑工作的神经连接路径,共享与人脑结构和功能有关的数据。

③“生物医学信息学研究网络”(BIRN)是一项旨在通过数据共享与合作促进生物医学研究的国家计划,为研究人员提供用户驱动的、基于软件的共享数据的框架。

7) 国家科学基金会

①“面向 21 世纪科学与工程的网络基础设施框架”(CIF21)将开发一系列整合的、可互操作的数据和工具,以支持科学与教育。

②“数据挖掘挑战赛”旨在解决大数据对人文和社会科学研究的挑战,利用新的研究方法检索、分析和理解数字图书、报纸、网络搜索结果、传感器、电话记录等数据集。

③“计算探险”项目为加州大学伯克利分校的研究人员提供了经费,以处理大数据研究挑战。分析学的根本性革新、云与集群计算及众包中伸缩性资源的系统构架、人类活动和智能的结合将

为当今自动数据分析技术无法解决的问题提供解决方案。

④“信息集成和信息学”项目旨在解决将传统科学研究数据转移到超大型异构数据库时所涉及的挑战和升级难题,如新型数据模型和表征方式的集成,以及与数据通路、信息生命周期管理和新平台相关的问题。

⑤开放科学网格(OSG)将全世界 8,000 多名科学家联合起来,共同探索搜寻希格斯玻色子等研究。

⑥理论与计算天体物理学网络(TCAN)项目旨在促进解释海量天体物理学数据的理论与计算方法。

8) 国家安全局 (NSA)

①“Vigilant Net”项目将开发保护计算机网络的数据可视化技术,从而促进和测试网络保护位置感知能力。

②信息安全与大数据的结合:美国情报机构(IC)开展了一系列协调、宣传和计划活动,与遍布美国政府、学术界和产业界的人士合作,使参与非保密工作的科学团体也可以了解情报机构的观点。

(来源于:《咨询参考》2012年第11期,检索:姜禾、张娟、刘晓,编译:姜禾、张娟,编译自 http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_press_release_final_2.pdf http://www.nsf.gov/pubs/2012/nsf12499/nsf12499.htm?WT.mc_id=USNSF_25&WT.mc_ev=click http://www.whitehouse.gov/sites/default/files/microsites/ostp/big_data_fact_sheet_final.pdf

麦肯锡发布《大数据：下一个创新、竞争和生产率的前沿》

麦肯锡环球研究院于2011年5月发布了《大数据：创新、竞争和生产力的下一个前沿》报告。该报告系统阐述了大数据概念，详细列举了大数据的核心技术，深入分析了大数据在不同行业的应用，明确提出了政府和企业决策者应对大数据发展的策略。麦肯锡认为只要给予适当的政策支持，“大数据”将促进生产力增长并推动创新。

1. 大数据及其作用概览

麦肯锡认为，“大数据”是指其大小超出了典型数据库软件的采集、储存、管理和分析等能力的数据集。该定义有两方面内涵：一是符合大数据标准的数据集大小是变化的，会随着时间推移、技术进步而增长；二是不同部门符合大数据标准的数据集大小会存在差别。目前，大数据的一般范围是从几个TB到数个PB（数千TB）。

大数据可以发挥重要的经济作用，不但有利于私人商业活动，也有利于国民经济和民生。数据可以为世界经济创造重要价值，提高企业和公共部门的生产率和竞争力，并为消费者创造大量的经济剩余。麦肯锡研究了“大数据”中尚未开发的巨大价值。例如，充分利用“大数据”的零售商将能够将营业利润率提高60%以上。“大数据”在公共领域也有极大潜力可挖。如果美国医疗保健行业有效利用“大数据”，就能把成本降低8%左右，从而每年创造出3,000多亿美元的产值。在欧洲发达国家，如果政府利用“大数据”提高运作效率，那么将节省至少1,000亿欧元（约合1490亿美元）的成本。而利用个人位置数据（personal location data）提供的服务将可以创造6,000亿美元的消费者经济剩余。

2. 大数据的关键技术

1) 可用于大数据分析的关键技术

可用于大数据分析的关键技术源于统计学

和计算机科学等多个学科，其中关于分析新数据集方法的研究仍在继续。需要注意的是，A/B测试、回归分析等技术也可应用于小数据集分析。

麦肯锡认为，可用于大数据分析的关键技术主要包括A/B测试、关联规则挖掘、分类、数据聚类、众包、数据融合和集成、数据挖掘、集成学习、遗传算法、机器学习、自然语言处理、神经网络、神经分析、优化、模式识别、预测模型、回归、情绪分析、信号处理、空间分析、统计、监督式学习、无监督式学习、模拟、时间序列分析、时间序列预测模型、可视化技术等。

2) 专门用于处理大数据的关键技术

麦肯锡认为，可专门用于整合、处理、管理和分析大数据的关键技术主要包括Big Table、商业智能、云计算、Cassandra、数据仓库、数据集市、分布式系统、Dynamo、GFS、Hadoop、HBase、MapReduce、Mashup、元数据、非关系型数据库、关系型数据库、R语言、结构化数据、非结构化数据、半结构化数据、SQL、流处理等。

3) 可视化技术

麦肯锡认为，可视化技术是大数据应用的重点之一，目前主要包括标签云、Clustergram、历史流、空间信息流等技术和应用。

3. 利用大数据的方法

麦肯锡对医疗保健、零售、公共领域、制造、个人位置数据等领域进行了重点分析，提出了利用“大数据”的五种方法：

1) 以时效性更高的方式向用户提供“大数据”。在公共领域，跨部门提供“大数据”能大幅减少检索与处理时间。在制造业，集成来自研发、工程、制造单元的数据可以实现并行工程，缩短产品投放市场的时间。

2) 通过开展数据分析和实验寻找变化因素

并改善产品性能。由于越来越多的交易数据都以数字形式存在,各机构可以收集有关产品或用户的更加精确和详尽的数据。

3)区分用户群,提供个性化服务。“大数据”能帮助机构对用户群进行更加细化的区分,并针对用户的不同需求提供更加个性化的服务。这是营销和危机管理方面常用的方法,但也可以为公共领域等带来变革。

4)利用自动化算法支持或替代人工决策。复杂分析能极大改善决策效果,降低风险,并挖掘出其他方法无法发现的宝贵信息。此类复杂分析可用于税务机构、零售商等。

5)商业模式、产品与服务创新。制造商正在利用产品使用过程中获得的数据来改善下一代产品开发,以及提供创新性售后服务。实时位置数据的兴起带来了一系列基于位置的移动服务,例如导航和人物跟踪。

4. 政策制定者应对大数据的策略

大数据不仅是企业竞争和增长的引擎,而且对于提高发达国家和发展中国家的生产率、创新能力和整体竞争力都有着重要作用。政策制定者需要认识到利用大数据可以刺激经济的下一波增长。为帮助企业获取大数据收益,政策制定者需要从如下6方面制定相关政策:

1) 加强大数据人才培养

政府可采取多种措施增加大数据相关人才供给,如实施教育培养计划、消除从其它国家和地区引进人才的障碍等。除此之外,政府应创造激励措施并对企业管理者进行数据分析技术培训。

2) 制定奖励措施促进数据共享

政府在创造数据共享和交易的有效市场方面可以发挥重要作用,包括制定知识产权方面规则、制定鼓励数据共享的奖励措施、强制要求收集并公开国企财务数据、以及面向公众开放和共

享政府部门活动和项目信息等。

3) 制定平衡数据使用与数据安全保护政策
政策制定者需要加强制定并执行关于商业和个人数据隐私的方针和法律,并通过强大的法律阻挡黑客和其它袭击。当然,政府、非营利组织和私人部门需要开发教育项目,以便公众理解哪些个人信息是可以获取的,如何使用、怎样使用,以及个人是否允许这种使用。

4) 建立有效的促进创新的知识产权框架

在大数据时代,数据价值链中的创新将不断出现,更好地产生和获取数据的技术也将出现。这些创新需要建立有效的知识产权保护体系,促进数据创造价值、数据共享和整合。

5) 克服技术障碍并加速关键技术研发

政策制定者加强制定IT工具或数据资源池的标准和指南,鼓励存在缺口的重要领域关键技术研发,推动行业标准制定机构制定覆盖IT工具和数据类型的标准,并给予资金支持、税收支出和减免、金融支持等激励支持大数据研究。

6) 确保信息通信技术基础设施投资

政策制定者应该使基础设施成为大数据发展的重要组成部分。很多国家对扩建基础设施制定了专门的激励措施。例如,美国政府出台了一系列货币奖励措施,鼓励宽带建设(如农村宽带工程)和实施电子医疗记录。

总之,政策制定者在人才、研发、基础设施和培育创新等关键领域能发挥重要作用,促进企业从大数据中获取最大收益。但是保持企业和公众间权利的平衡是个艰巨的任务,政府在赋予企业在更大范围使用数据以获取潜在收益同时,要减轻公众对隐私和个人信息安全的担忧。

(来源于:《资讯参考》2012第11期,检索:唐川、刘晓,编译:田倩飞,编译自
http://www.mckinsey.com/mgi/publications/big_data/pdfs/MGI_big_data_full_report.pdf)

国家数学与交叉科学中心科研发展研讨会暨工作总结会召开

2013年1月24日，中科院国家数学与交叉科学中心（以下简称“交叉中心”）科研发展研讨会暨年终总结会在北京召开。交叉中心执委会成员、各研究所负责人及中心青年科研人员代表等30余人出席了会议。

会议由交叉中心主任郭雷院士主持。交叉中心副主任、数学院执行院长王跃飞研究员首先传达了2013年中科院工作会议精神，介绍了科学院2013年总体工作思路，即按“创新2020”跨越发展体系总体部署，以高科技产出质量和贡献为中心，深入实施“一三五”规划，抓好重大成果产出，深化科技评价改革，加快推进一流管理，深化拓展国际合作，实现“三位一体”中科院建设的新跨越。

郭雷院士指出，此次交叉中心科研发展研讨会的主要目的是集思广益，充分探讨如何利用好现有条件开展好交叉中心的科研工作。交叉中心具备良好的综合交叉实力和基础，国家和科学院都对交叉中心的发展寄予厚望。面对国内外挑战，如何保持自身特色，原文忠志，紧扣国家战略需求，开创新的学术特色，用自身实力和重大研究成果向国家和科学院交一份满意的答卷是交叉中心的使命和责任，需要认真努力，稳步推进。

随后，交叉中心副主任何桂英研究员就2013年中心在人才、经费、学术交流等方面的相关计划和安排向与会者做了介绍，并征求了大家意见。

研讨会上与会者纷纷发言，肯定了交叉中心成立以来专注交叉领域所组织的科研工作，讨论了未来进一步发展的一些计划，同时也对交叉中心的工作提出了中肯的建设性意见。

会议还听取了6位中心青年科研人员代表的研究工作总结报告。与会人员充分肯定了青年科研人员进入中心以来所取得的成就，并就他们报告中存在的问题以及他们在研究和生活中所遇到的问题进行了交流探讨。

此次会议不仅是对交叉中心过去一年的工作进行了总结，也是新的一年工作的一个良好开端。大家表示，在新的一年里，要努力开拓创新，团结协作，把交叉中心的工作做得更好。



中国科学院国家数学与交叉科学中心标志

为全面做好中心的各项工作，增强广大职工和社会各界对交叉中心的认识和了解，自即日起，正式启用中国科学院国家数学与交叉科学中心标志。

图标由文字和图案两部分组成，“NCMIS”为中国科学院国家数学与交叉科学中心英文名称的缩写，图案以Mathematics单词首字母的融合与交叉构成，宛如绽放的花朵，充满魅力与灿烂之花，寓意交叉中心是数学与多个学科的交叉融合，象征着中心成立之初由6个研究院及1个分中心组成，彰显数学与其他学科交叉研究、协同合作的广阔前景与未来。图标也体现了交叉中心对多学科研究的重视、包容和兼容，以及尊重科学、崇尚科学的作风。

图标分为彩色版和单色版两种，彩色版是由7种颜色构成，单色版以紫色为基本配色。中国是世界上最早人工合成紫色颜料的国家，紫色，高贵典雅，代表高洁和精神，是交叉中心视觉形象标准色。此外，图标也融入北京元素，表明了交叉中心服务社会、奉献社会的责任和使命意识。在具体使用过程中，为方便起见，单色版可以采用其他颜色。

图标用于交叉中心的各类活动、宣传出版物、纪念品、印刷品等有关制品上。交叉中心所属各研究所、分中心举办各类活动时使用。未经许可，不得用于经营及商业活动。交叉中心办公室负责图标的授权使用管理。

