



北京航空航天大学
BEIHANG UNIVERSITY

《学术报告》总结

学 院	<u>计算机</u>
姓 名	<u>陈志伟</u>
学 号	<u>SY1406108</u>
导 师	<u>马殿富</u>

北京航空航天大学研究生院

2016 年 09 月 12 日

目 录

Model Driven Engineering, Functional Safety, and High Tech Systems	3
Synchronization, Coherence and Event Ordering in Multiprocessors	4
基于 WiFi 指纹的室内定位技术.....	5
面对面为你揭秘大型软件开发.....	6
百度搜索引擎技术以及阿拉丁和搜索推荐技术.....	7
游戏开发技术揭秘.....	8
IT 技术人才必备神器——性能测试.....	9
“二十一世纪的计算”学术研讨会	10
大数据领域的三个问题.....	11
Deep Learning with GPUs - Technologies from NVIDIA	12

Model Driven Engineering, Functional Safety, and High Tech Systems

这次报告的作者是荷兰埃因霍芬理工大学的 Mark van den Brand 教授，他主要的研究方向是软件工程、编译器和编程语言等，和我目前的研究领域非常相关。

MVD Brand 教授主要介绍了领域专用语言(domain specific language / DSL)，其基本思想是“求专不求全”，不像通用目的语言那样目标范围涵盖一切软件问题，而是专门针对某一特定问题的计算机语言。几乎自计算机发明伊始，人们就开始谈论 DSL、使用 DSL 了。

特定领域语言(DSL)的语义的形式定义是使用 DSL 进行模型转换正确性验证的一个关键前提。Brand 教授介绍了他们为系统的并发、连通对象(SLCO)的规格实现了一个 DSL 的语义原型。使用这种原型，DSL 所指定的模型可以转化为标记过渡系统(LTS)。这种把模型变换为 LTSs 的方法使我们能够轻易的应用现有的工具进行模型的可视化和验证。

最后，Brand 教授还介绍了 DSL 语言应用的优点，主要是：(1) 高级别的重用：如果仅适用通用编程语言，则每次只能解决一个问题，但如果应用特定领域开发方法设计并实现一些特殊语言，每个特殊语言可以高效地解决一类相似的问题；(2) 使用 DSL 的软件架构可以跨接软件工程过程各阶段之间的鸿沟，特别是通过代码生成可以很好的进行设计和实现阶段的衔接；(3) 使用 DSL 表达的模型，可以在问题空间这个较高的抽象层次进行验证，这意味着可以在开发周期的更早期发现因为理解和表述而造成的错误；(4) 一个模型中具备了重要的业务知识，将解决方案从一种技术迁移到另一种技术，或在同一技术的不同版本之间迁移，就变的相对容易。一般通过适当修改生成器或解释器就可以做到。

通过这次报告让我感受到了 Brand 教授高深的学术造诣，同时也使我深深的体会到 DSL 语言在形式验证领域的强大之处，它可以很好的应用到目前实验室的研究领域。但在由于 DSL 语言的专用性，在我们广泛使用它之前，还需要定义一套语义原型，这需要结合具体的问题领域进行进一步的研究。总之，DSL 语言为我们的研究指明了另一条路。

Synchronization, Coherence and Event Ordering in Multiprocessors

随着单核处理器的主频接近 4GHz 的极限，单个处理器性能的提升已经不能再依靠不断提高的时钟频率，因为功耗会急剧上升，产生大量的热。因此如今为了继续提升处理器的性能，多核处理器便应运而生。

Michel Dubois 教授在这次报告中为我们讲述了多核处理器中存在的同步、一致性和事件有序性问题。这个问题他早在 1988 年 2 月就在 IEEE 计算机上发表的一篇文章中提到，从那时起，它已成为并保持在计算机系统结构，甚至计算理论的主要问题之一。理解这个问题和它的解决方案对设计正确的共享内存系统和 CMPs 来说非常重要。然而，随着时间的推移，Dubois 教授仍然觉得今天，包括在那之后的所有时间，所有已写或发表的论文，所有的书，所有的专题讨论会和研讨会，在内存连贯性的讨论和一致性这个问题上，学生和研究人员都一知半解，因为这确实是一个非常困难的课题。

讲座中 Dubois 教授首先介绍了同步机制和基本的一致性协议，这些机制的参数将被降低到严格的最小值以简化逻辑参数。接着介绍了一致性的两种视角，即存储的原子性和普通的一致性，但这两种视角都是错误的及错误的原因。然后介绍了多核处理器的内存一致性模型，包括顺序处理器的语境定义和乱序处理器中可能违反内存一致性的情况。最后，Dubois 教授谈了谈他对一致性的看法，我们是否真的需要一致性？这个问题值得我们在这个多核处理器如火如荼的时代冷静下来深究。

Dubois 教授以在一致性、同步和共享存储的多处理器的内存访问次序的贡献而闻名于世，他是一位真正的大家，能入木三分的思索一些最本质的问题，他的科学的思维方式和精神值得我认真学习。现如今我们开发的各种激动人心的应用，研发的各种智能的硬件等，这些东西确实给我们的生活带来了方便，甚至成为了我们日常生活不可缺少的一部分，但我们同样不能忘记这样一群人，他们关注着科学技术背后最本质的东西，保持着初心，他们的研究是现代技术的基石，更值得我们永远铭记。

基于 WiFi 指纹的室内定位技术

这次报告的作者是学校的牛建伟教授，他主要的研究方向是嵌入式与移动计算、数字图像分析与处理。

牛教授首先介绍了一些室内定位技术出现的背景。随着近代通信技术的发展，基于位置服务的应用越来越多。不同商业应用的增加促进了基于位置服务的发展。由于多径效应的存在以及信号的过度衰减，使得传统的基于 GPS 定位的技术在室内无法使用。因此，大量的室内定位技术的应用应运而生。而越来越多公共场所都部署了大量的无线局域网设备，因此可以利用 WiFi 信号来进行室内定位。

接着介绍了当前 WiFi 定位技术中所存在的问题。虽然 WiFi 设备被大量部署应用，使用 WiFi 指纹进行室内定位也越来越方便，但传统定位方法中存在功耗较大，定位不精准等问题，因此研究低功耗接口探测 WiFi AP (Access Point) 信号生成位置指纹技术具有重要实用价值。

然后，牛教授介绍了他所提出的室内定位新方法。该方法使用低功耗的 ZigBee 接口探测 WiFi 信号生成位置指纹，然后结合 Mapper 提供的 AP 信息，使用对齐算法形成 (BSSID, RSS) 向量，最后通过对该向量进行分类来确定移动设备的位置。与使用 WiFi 接口的室内定位方法相比，该方法在能够取得相似的定位精度的条件下可以节省约 60% 的能量，比较适用于能量有限的移动设备，且不需要改变已有的网络设施或是添加新的硬件。

通过这次报告我学习到了 GPS 的局限性，以及人们为了克服 GPS 无法定位室内所作出的一些努力。同时，牛教授作为我们 ACT 实验室的老师，是一位真正做学问的好老师，近年来的研究成果颇丰，值得我们每个人认真学习，不经一番彻骨寒，怎得梅花扑鼻香。

面对面为你揭秘大型软件开发

这次报告其实是华为公司到学校做的一次宣讲，也是为了庆祝北航华为创新俱乐部的成立。俱乐部旨在加强北航在校生与华为公司的校企对接和交流，为学生和企业提供沟通的平台，给同学们提供更多近距离接触华为的机会，也为华为推荐更多优秀的北航人。

报告中，华为的首席软件架构师杨宏杰描述了他们团队进行的大型电信软件的开发过程，与互联网行业的软件快速迭代和快速适应调整不同，他们要求开发出的软件要达到电信级的安全、稳定性，所以会有更加严格的测试。实际中，一个 bug，华为的芯片得重新流片；一个 bug，华为的基站得退服，客户得跟政府解释，所以互联网公司追求快，华为追求稳。

报告结束后，进入了自由提问的环节，有 4 位来自华为不同部门的北航师兄负责解答我们的提问。在所提出的问题中，大家最关注的一个焦点就是华为的加班文化和工作压力的问题，这个问题也代表了这么多年里外界对华为的整体印象，因为近几年来不断爆出华为员工过劳死的现象。师兄们解释到，在华为加班确实是一种文化，但他们所在的部门还好，可以准点下班。同时，他们也提到了自己的一些看法，在人生最好的时代，你是选择好好奋斗一场，还是选择安逸。“Equal pay for equal work”，华为的高压力是和工资相匹配的。

这次报告改变了我长久以来对于华为公司的看法，站在他们的角度来想想，华为公司从一个民营企业发展成为世界 500 强公司，靠的是什么，就是公司的狼性文化。大鱼吃小鱼，快鱼吃大鱼。没有快速的产品迭代和一流的服务，华为公司肯定走不到今天的地步，更不可能在国际电信市场上打败思科、爱立信等巨头。未来是否会选择华为，还是看每个人的需要吧，一个关键的选择也决定了以后一生的道路。总结一下，华为要的是最勤劳的人，而不是最聪明的人。

百度搜索引擎技术以及阿拉丁和搜索推荐技术

这次报告的作者是百度主任工程师黄际洲先生，他主要负责百度搜索推荐。自加入百度以来，曾先后负责过百度计算需求分析、图片信息库及其创新应用、图片搜索交互与引导等技术方向，并在这些方向上进行了多项重大创新，其成果已应用于百度网页搜索及图片搜索等产品，累计申请专利 70 余项，是一个技术牛人。

报告中黄先生首先介绍了百度搜索引擎技术，包括了百度搜索引擎技术的发展，以及在移动互联网时代百度搜索引擎技术的创新、发展和突破。接着介绍了阿拉丁和搜索推荐技术。阿拉丁是搜索引擎百度公司推出的一个通用开放平台，它将接口开放给独特信息数据的拥有者，从而解决现有搜索引擎无法抓取和检索的暗网信息。

很多人类信息并没有被 Web 化，所以目前能被搜索引擎检索到的信息只占人类所有信息中非常小的一部分，更多的信息都存在于 Hidden Web，也就是“暗网”之中。这是一个基于百度网页搜索的开放的数据分享平台，广大站长和开发者可以直接提交结构化的数据到百度搜索引擎中，实现更强大、更丰富的应用，使用户获得更好的搜索体验，并获得更多有价值的流量。总之，阿拉丁与搜索推荐是百度搜索引擎的核心组成部分，在用户需求的精确高效满足以及用户引导方面发挥着无可替代的作用。最后，黄先生还分享百了度在需求满足及需求引导方面的技术进展与成果，以及我们所进行的一些创新探索。

通过这次讲座极大的扩宽了我的视野，我们所到的网络只是整个互联网的极小的一部分，隐藏之下的是看不见的一望无际的暗网，就像宇宙中的暗物质一样广阔。百度公司的阿拉丁计划意义重大，但也有一定的局限性，不能完全靠人来添加未知的网页，要真正解决暗网的问题，还是需要一代又一代人努力的探索，力争在搜索引擎技术上取得新的突破性成就。

游戏开发技术揭秘

本次报告的作者是网易游戏天下事业部技术经理陈伊力先生，他是网易出品的著名 RPG 游戏“乱斗西游”的主程。网易在线游戏在中国 MMORPG 游戏市场保持领先地位，是网络游戏自主开发和成功运营的大成者，在开发实力、产品线和市场占有率上成为国产网游的领军者。

报告中陈先生首先介绍了不同游戏开发侧重的分类和游戏开发的基本流程。游戏类型的提出是游戏作为一种娱乐形式发展到一定成熟度的标志。游戏按类型可分为角色扮演、第一人称设计、模拟经营、RPG 等。游戏的分类是一种习惯和约定俗成，主要用于业界人士以及玩家之间的交流沟通。游戏设计与开发主要有三大基石：策划是在现有条件、技术基础上规划、策划展现在玩家面前的情景；技术主要用到的是三维图形和人工智能技术；过程要求控制稳定性、设计周期和设计成本。

接着介绍了游戏过程中需要用到的关键技术，包括网易自研的 NeoX 游戏引擎。最后是对未来游戏的发展方向的畅谈，未来更多端游开发商将进入手游市场，端游改编手游、跨平台游戏将成为趋势；细分领域游戏，独立游戏将有更大的空间；VR 将是 2016 的大热门，但很难在玩家中爆发，主要问题是虚拟现实如今面临着诸多问题，比如缺乏技术标准，内容匮乏，企业良莠不齐等诸多问题。总的来讲，VR 是游戏的未来，但不是现在。

通过这次讲座我学到了游戏开发方面的一些知识，同时也不经为网易这个中国非常有创新精神的公司鼓掌。在当年国内游戏市场被国外产品霸占时，是网易通过艰苦卓绝的努力开发除了两款至今仍长盛不衰的网游：大话西游和梦幻西游，守住了国内游戏市场最后一丝尊严。最后，希望网易能再接再厉为广大用户开发出更多更好的产品，有朝一日能成为 BAT 的另一极。

IT 技术人才必备神器——性能测试

本次报告的作者是章张先生，他是一位阿里云技术专家。阿里云创立于 2009 年，是中国的云计算平台，服务范围覆盖全球 200 多个国家和地区。阿里云致力于为企业、政府等组织机构，提供最安全、可靠的计算和数据处理能力，让计算成为普惠科技和公共服务，为万物互联的 DT 世界提供源源不断的新能源。

报告中章张先生主要介绍了阿里云性能测试服务。阿里云计算发布的性能测试服务 PTS（Performance Test Service），可以帮助开发者通过分布式并发压力测试，模拟指定区域和指定数量的用户同时访问，提前预知网站承载力。利用该服务，开发者一方面可以找到自身网站或 APP 系统性能的瓶颈从而进行优化，另一方面可以充分了解系统性能指标，便于未来扩容。

性能测试服务是集测试机管理、测试脚本管理、测试场景管理、测试任务管理、测试结果管理为一体的性能云测试平台。针对阿里云用户复杂的分布式应用，PTS 可以快速扩容，动态配置域名，满足不断增长的集群压测需求。简而言之，完全基于阿里云平台构建的性能测试服务，能够针对性地提供多种模拟真实场景的施压方式，以帮助阿里云计算平台上的开发者获得精准的测试结果。同时，提供完善的性能结果报表和强大的性能分析日志，便于开发者能够有针对性地进行系统的调整优化。

最后，章张先生还畅谈了当前的就业形势，并针对毕业生与企业所需技能差距和就业方向与岗位解答了我们的一些疑惑。通过这次报告我看到了阿里云作为中国企业的在云计算领域追赶世界云技术巨头 AWS 的曙光，虽然阿里云暂时还没有 AWS 那么大的体量，但相比其他国内的公司，他是国内做的最好的。在最新的云计算排名中阿里云已经上升到了世界排名的第 3 位，这是非常了不起的成就。依靠着阿里人不断的技术创新，有朝一日阿里云定能跟云计算领域的巨无霸 AWS 一决雌雄。

“二十一世纪的计算”学术研讨会

第十七届“二十一世纪的计算”大型国际学术研讨会是在清华大学新清华学堂隆重举行的。本届大会邀请到两位新晋图灵奖获得者作为演讲嘉宾——微软研究院首席研究员 Leslie Lamport（2013 年图灵奖得主）和麻省理工学院客座教授 Michael Stonebraker（2014 年图灵奖得主），与中国高校师生分享其各自在人工智能领域的研究成果和最新观点。除此之外，微软全球资深副总裁 Peter Lee、微软亚太研发集团主席兼微软亚洲研究院院长洪小文博士及其他多位全球顶尖的计算大师们也纷纷发表了主题演讲，集中展示了计算科学领域的最前沿技术及未来的发展趋势。

在报告中 Leslie Lamport 博士分享的题目是“用数学视角去看待编程”，他希望人们用数学的视角去看待编程，是数学教会了我们抽象思考，而 C/C++ 之类的编程语言无法教会我们抽象思考。在由 0 和 1 组成的数字系统中，任何状态之间的转变，都可以用数学方法去描述，从简单的一维时钟到二维时钟，甚至更复杂的系统，最终抽象出来的都是数学公式，并可以用数学方法去证明它。但对工程师来说，编程的时候无需证明这个过程本身，只需要写代码做出东西即可；但如果从数学的角度去看待，尽管它不能让程序运行更有效率，但能让程序更逻辑严谨。

在此基础上，Leslie 介绍了一个名叫 TLA+ 的语言，它是一种建模语言，可以用于将通信协议进行建模检测，以确认能否实现安全目标，抵制攻击。截至目前，TLA+ 已经在亚马逊、英特尔和微软得到了广泛的使用，帮助工程师们发现了编程中的很多错误。Leslie 表示，“TLA+ 能够教会我们用数学的方式去进行抽象思考。对于工程师来说，需要把复杂的系统尽可能简单实现，描述这些系统的数学原理非常简单，但编程语言就相对复杂。为了满足执行效率等其他因素的需要，纷繁复杂的编程语言无法让工程师更好地去思考。但现在人们却反过来，觉得 C/C++ 是非常简单的语言，而数学是极其复杂的东西，这完全是本末倒置。”

通过 Leslie Lamport 博士的报告，我更加感受到了数学在编程领域的重要性。我们以前只是把数学当做编程中的一个工具帮助我们去建模，去测试一些用例而非用数学来思考，现在是时候进行一些转变了。随着编程语言的发展，我们需要从以前的编程思维，即按照计算机执行那样去思考，转变为数学思维，即使用数学表达式编程，让编程语言去处理计算机底层相关的事情，使人真正的解放出来。我也相信这是未来的编程语言的发展方向。

大数据领域的三个问题

本次报告的作者是 Michael Stonebraker 教授，他是数据库领域的顶级专家、美国工程院院士、冯诺依曼奖的获得者、2014 年图灵奖得主。他发明了几乎所有现代数据库系统所用的概念，创办了无数成功的数据库技术公司。

报告中，Michael 教授与大家分享了他所理解的“大数据到底意味着什么”，以及该领域的诸多产品和技术问题。在 Michael 看来，大数据无非就是要解决好三件事：一，处理大量数据（Volume）；二，很快的处理速度（Velocity）；三，解决好数据来源繁杂的问题（Variety）。

对于大数据的数据量太大这个问题，Michael 表示，市面上的很多数据仓库产品已经可以很好地应对这个问题了，例如采用行存储、行执行方式的 Netezza、Oracle，采用列存储、行执行的 Teradata/Asterdata 和 EMC/Greenplum，以及采用列存储、列执行的 HP/Vertica、Sybase/IQ、Paracel/RedShift、DB2/BLU 和 Microsoft PDW。列存储的速度比行存储方式快了 50 倍，所以他更喜欢的一种解决方案是：为什么不可以有一个阵列模式的数据系统？在他看来，阵列数据库将会是未来的赢家。

在大数据的处理速度方面，Michael 认为，NoSQL 采用全新架构，速度更快。不过如今 NoSQL 和 SQL 越来越有融合的趋势，可以将融合之后的 OLTP 称为 New SQL，这也是他所青睐的。

而对于数据来源的多样化，尤其是在面对成千上万个数据源时，传统的先进行数据提取、再去收集并建立全景模式的方式已经不够用了，但传统供应商也无法解决这个问题。Michael 介绍了一家初创公司，他们在可扩展性方面做得相当不错，其原因在于该公司借助机器学习来处理多样的数据源以及巨大的数据整理工作，效率因此提升了许多。

通过这报告我领略到了学术大师的风采，同时对当前如火如荼的大数据有了一个基本的认识。同时，Michael 教授介绍的大数据领域最新的进展，极大的开拓了我们的视野。大数据是未来的发展方向，我们将进入大数据时代，每个人都应该具备一定的大数据方面的知识。

Deep Learning with GPUs - Technologies from NVIDIA

本次报告的作者是 Junjie Lai 博士，他的主要研究方向包括机器学习，深度学习以及计算机视觉等方面。现在 Lai 博士是 NVIDIA china 高性能计算组的技术总管。除了领导这个团队，他还与来自其他知名互联网企业的开发者共同研究如何用 GPU 加速深度学习应用。

报告中 Lai 博士主要介绍了 GPU 在现在火热的深度学习领域的应用。GPU 是图形处理器的缩写，又称显示核心、视觉处理器、显示芯片，是一种专门在个人电脑、工作站、游戏机和一些移动设备上图像运算工作的微处理器。GPU 传统上主要用于计算机图像发面的应用，在引入 CUDA（统一计算设备架构）的图形处理器后则已经在高性能计算领域显示出巨大的潜力，并被广泛地应用在不同的领域。

近年来，深度学习研究取得重大突破。深度学习是机器学习领域中对模式（声音、图像等等）进行建模的一种方法，它也是一种基于统计的概率模型。深度学习的学名又叫深层神经网络（Deep Neural Networks），是从很久以前的人工神经网络（Artificial Neural Networks）模型发展而来。这种模型一般采用计算机科学中的图模型来直观的表达，而深度学习的“深度”便指的是图模型的层数以及每一层的节点数量，相对于之前的神经网络而言，有了很大程度的提升。

深层神经网络参数的训练通常消耗大量的计算能力，越来越多的研究者开始使用 GPU 来加速他们的深度学习研究，Lai 博士分享了 NVIDIA 在这一方面的进展。除了提供硬件加速，NVIDIA 提供 CUDA 开发工具、性能分析工具和大量的数学库。特别地，cuDNN 库是 NVIDIA 专为深层神经网络计算而开发的。所有这些努力都使得使用 GPU 深度学习研发更轻松、更高效。

通过这次报告使我学习到了 GPU 方面知识，同时既然 GPU 凭借其强大的计算能力可以用在深度学习方面，那也可以把其应用到其他方面，作为备用的第二个 CPU 来使用。当前机器学习正火，突然感觉有点遗憾自己错过了这一浪潮，但未来若有机会，还是会往这一方向转移。最后，感谢 Lai 博士的精彩报告！