

谈“计算机学科基础研究”

包云岗 2008-6-29

(注: 4月24日, 计算所领导与智能中心员工学生面谈讨论, 我做了题为“体系结构基础研究”的报告, 表达了对计算机学科基础研究的观点, 引起了大家对该话题的讨论。大家认为基础研究要重视有效产出、要重视资源利用效率。本文是在上述报告及大家讨论基础上的进一步总结与思考, 与大家一起探讨。)

长期以来, 计算所以完成国家大项目为主要任务, 近几年来, 所里开始重视基础研究。计算所的学术论文数量和质量都有明显的提高。与此同时, 兄弟院校(如清华等)在信息学科方面也突飞猛进, 屡有高水平文章发表于国际顶级会议期刊, 大陆的计算机学科基础研究能力不断增强。另一方面, 由于我们人均科研资源有限, 中国的基础研究更应该重视有效产出, 而不能简单地学习国外的“Paper-To-Paper”模式。这些讨论引起了我对计算机学科基础研究自身特点的好奇, 于是便展开了对计算机学科基础研究的一些调研。

历史地看待计算机学科基础研究

自第一台计算机诞生以来, 计算机逐渐发展为一门独立的学科。而到1960年代, 随着越来越多的大学与企业开展计算机技术的研究与推广, 计算机也逐渐地从完全为政府、军事等服务走向了民用应用领域, 开拓了新的产业。进了21世纪, 信息技术甚至形成了高达几万亿美元产业规模, 更是各国财政收入的重要来源之一, 因此各国政府对信息产业发展都非常重视。

2003年, 美国科学院的国家研究顾问团(National Research Council of the National Academies)对信息产业的历史发展进行了调研、分析与评估, 形成了供美国政府决策参考的重要报告——《信息产业创新(Innovation in Information Technology)》。这份综合报告旨在强调基础研究对信息产业的作用, 同时揭示了政府、大学、企业在基础研究与产业推广中的重要作用与互动关系。

回顾历史, 上世纪60年代信息产业脱离政府控制、独立发展并步入成熟期。在60年代初, 一些企业开始建立各自的研究所, IBM于1961年成立了托马斯·华生实验室(Thomas J. Watson Research Center), 主要进行基础研究以及应用研究。同时, IBM的计算机也开始在市场上获利, 如1964年推出的System/360就是计算机进入商业计算的重要标志。企业的基础应用研究使得政府可以更加专注于支持一些更长期、更具风险性、更基础的研究, 另一方面也吸引了更多的人选择信息技术作为自己的职业生涯规划。这个阶段, 一些重要的大学都开始设置一门新的学科——计算机科学(Computer Science), 如1965年Stanford与CMU成立了计算机系, 而MIT也于1968年成立了计算机系。因此, 上世纪60年代正是信息领域繁荣的起点, 也是计算机学科基础研究活跃的起点。

以1960年代为时间起点, 《Innovation in Information Technology》报告中列出了19个重要IT技术的发展图谱¹(见图1与图2)。这19个重要技术现在的市场规模均已达到了10亿美元甚至更多。直观上, 我们可以看到错综复杂的箭头。仔细分析, 发现这些箭头可分为以下几类:

- **局限于一个技术内大学与企业之间的箭头:** 这些箭头体现了大学基础研究与企业R&D的相互推动作用。例如, IBM最早提出了关系数据库(Relational Database)概念, 但是未能将该技术商业化。Berkeley等大学在美国NSF资助下对IBM提出的概念展开了进一步研究, 并将该技术商业化, 孵化了一些start-up公司。而企业在发展的过程中遇

¹ 报告指出, 所列出的19个例子并不能完全覆盖信息产业的所有技术, 但具有很强的代表性。

到新的技术问题又会反馈到大学进行基础研究。这种“大学—企业”互动模式产生了积极效果，推动了许多技术的商业化。

- **科研实体与产品、市场之间的箭头：**科研实体（包含大学与企业）在某项技术的市场化过程中都起到了重要作用。有趣的是，从图中可以看到，很多技术的第一个产品最先是由大学推出，而且在扩大市场规模中大学也起着不可或缺的作用。仅仅从这两个图来看，大学与企业对产业推动并没有明确的分工²。

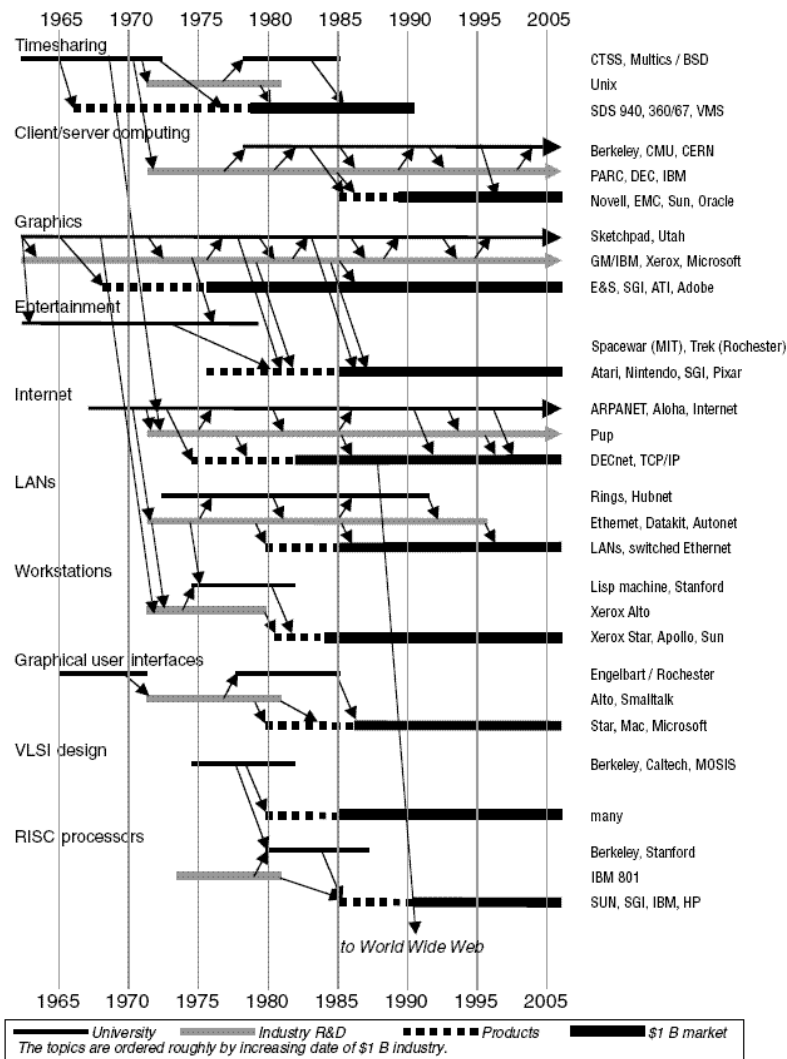


图 1. IT 技术的发展历史(1): 大学基础研究 + 企业 R&D + 规模市场

技术的发展正是得益于其他技术的进步，多种技术的相互结合往往能成倍的扩大技术的影响力。如 RISC 思想最初由 IBM 提出后并没有产生很好的市场效应，直到 70 年代末 80 年代初 Berkeley 与 Stanford 开展 VLSI 研究后才有进一步发展。VLSI 设计的研究大大地推动了 Synopsys, Cadence, Mentor 等公司的电路自动化设计技术发展，形成了高达 30 亿美元的市场。而正是这些电路自动化技术成为了其他重要技术（如 RISC 处理器）商业化的基础使能技术。

从以上分析可以看到，虽然信息领域快速发展的背后存在错综复杂的关系，但实际上这些关系可以最终归结为两类实体（大学、企业）和两类活动（基础研究、产业化）。

计算机学科基础研究与产业化

在政府管理下，上述两类实体所从事的两类活动对信息产业发展起实质性推动作用。仔细观察它们之间的关系，我们可以更进一步地挖掘一些重要的结论：

² 这种现象的产生可能得益于美国大学完善的技术转移与企业孵化制度，例如有许多著名的企业（Sun、Cisco、Yahoo 等）都离不开斯坦福大学的支持。（甚至 Cisco 早期与斯坦福大学之间围绕知识产权问题产生纠纷时，斯坦福大学在知识产权方面依然采取开放的态度，并未深入追求。）

- 产业化离不开基础研究（大学基础研究与企业 R&D）；
- 大学基础研究与企业 R&D 互动反馈周期约 3~5 年；
- 从基础研究到第一代产品大约需要 5~10 年；
- 从基础研究到 10 亿美元产业规模大约需要 10~15 年；

几乎所有技术的产业化都有大学与企业共同推动（图中的带箭头的水平线）。其中，大学基础研究侧重于开放的关键点、难点等共性问题，而企业 R&D 则倾向于更容易使用专利保护、更容易商业化的应用技术研究。以语音识别（Speech Recognition）技术的发展为例³，Bell 实验室最

早于 1960 年代开始研究语音处理专用机，然而进展非常艰难。Bell 实验室的执行主任 J. R. Pierce 在《The Journal of the Acoustical Society of America》上撰文称，“……语音识别太遥远，对人们吸引力好比水变油、大海淘金……”幸运的是，在 CMU 的参与下，语音识别开始有了新的转机。1971 年到 1976 年，语音识别可以达到了 90% 以上的识别率。1980 年代统计方法提出使得语音识别得到了新的发展。1980 年代到 1990 年代，MIT、SRI 进一步推进语音识别技术。而如今，语音识别已经多媒体重要技术之一，广泛地应用在各个领域。

图 3 则是一些人机交互技术的从基础研究到产业化的时间表⁴。从图中可看到，许多我们日常工作中依赖的技术（如鼠标、Window 等）都是经过相当长的时间才从大学基础研究推向市场。

语音识别与人机交互技术发展的例子正是上

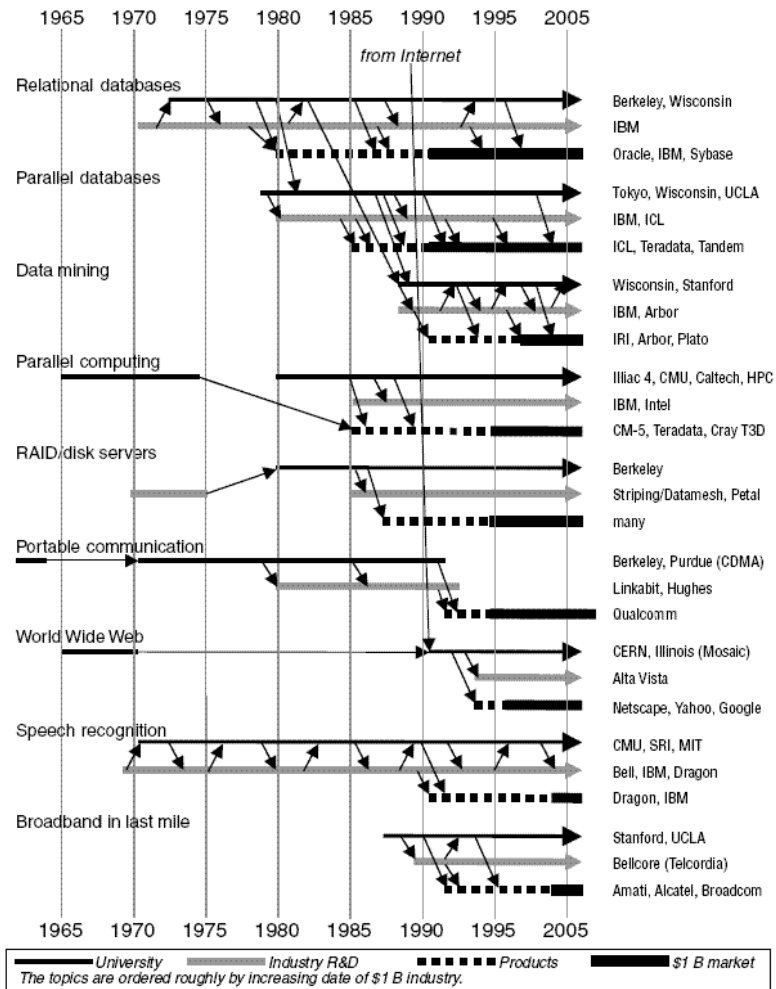


图 2. IT 技术的发展历史(2): 大学基础研究 + 企业 R&D + 规模市场

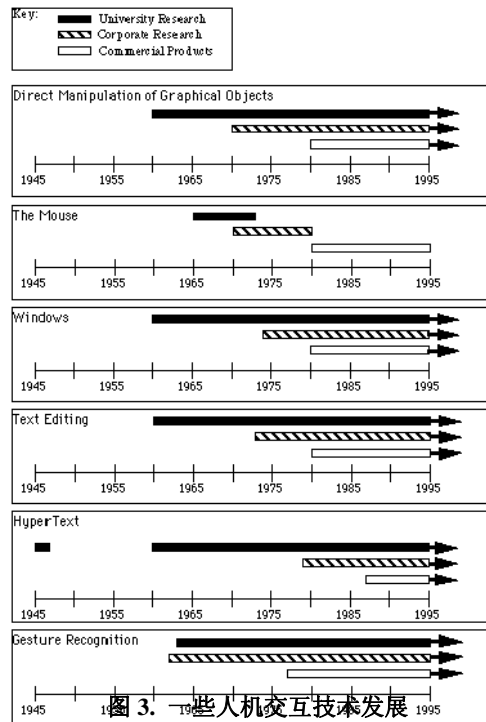


图 3. 一些人机交互技术发展

³ Automatic Speech Recognition: An Overview, Julia Hirschberg.

⁴ A Brief History of Human Computer Interaction Technology, Brad A. Myers.

述总结的 4 条结论的典型写照。因此，把握两类实体与两类活动的内在联系及其规律对我们认识信息领域基础研究有很重要的意义。

计算机学科基础研究的特点

所有学科的基础研究都具有一些普遍规律，但计算机学科也有一些内在的特点。总结来看，计算机学科基础研究有以下特点：

- **周期长：**从基础研究到初级产品一般需要 5~10 年。以现在大学及科研单位博士 5~6 年的培养周期来看，至少需要到 1~2 代无重叠周期博士参与到这些基础研究中。因此，基础研究在课题选择初期就需要有前瞻性，同时需要有稳定的领导人员、项目资金和合理的人才梯队来维持其研究的延续性。
- **实验工程型：**计算机学科是一门特殊的“人造学科”（Synthetic Discipline），研究的是人类自身创造的事物的现象与规律。由于是人造学科，研究人员可以很容易地创造新的计算现象或者新的实现方法。但不管创造了什么，最重要的是要证明比其他方法更好⁵。而且，计算机学科多是以工程项目开发的形式开展基础研究，需要科研人员有很强的项目管理意识。
- **非预期结果：**为了实现基础研究的既定目标，在研究过程中往往会产生一些副产品，即研究目标以外的非预期结果。这些非预期结果往往也很重要，也是产业化的技术来源之一。如 1960 年代的为方便大型机共享用户交流的 email、即时通信等技术在现代已经成为人们日常生活中的重要通信手段。
- **技术交叉发展：**计算机不同子领域的技术相互影响，一种技术的发展很可能会成为另一种技术发展的使能基础。因此，关注多种技术的发展在计算机学科的基础研究非常重要。
- **易产业化：**总体来看，计算机学科的基础研究比其他学科更容易商业化，这为从事计算机基础研究人员提供很多创造社会价值、经济价值的机会。因此，计算机学科基础研究的方向、课题选择也该考虑这些因素。

企业 R&D 主要侧重于应用技术研究，而基础研究主要集中在大学。因此，大学在计算机基础研究中扮演了重要的角色。图 4 是 Berkeley 的 David Patterson 等在报告《The Parallel Computing Landscape: A View from Berkeley 2.0》中总结的十三所大学对上述十九个重要技术产业化的贡献。与企业相比，大学在基础研究方面至少有以下两个重要优势：（1）大学可以同时进行多种技术的研究，

多学科的聚集使得大学更容易进行学科交叉研究、技术交叉研究；（2）大学研究具有更强的流动性、开放性，“大学—大学”之间、“大学—企业”之间都可以很方便的流动，而企业之间往往由于各自的商业利益在合作方面相对保守。当然大学也有其他的一些优势，比如可以进行长期的研究而不考虑利润，大学的教育与研究结合方式更有利于人才培养。

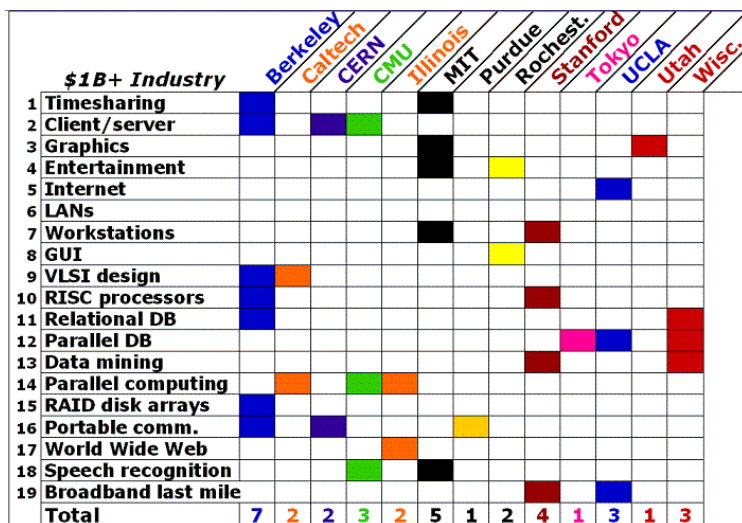


图 4. 十三所大学对十九个重要技术产业化的贡献

⁵ Academic Careers for Experimental Computer Scientists and Engineers, 1994, National Academy Press.

因此，一些顶尖大学（如 Berkeley, MIT, Stanford 等）正是利用自身的优势，并在与企业的互动中良性循环发展，即推动了计算机产业的发展，培养了杰出的计算机科研人员，提升了大学的学术影响力。

基础研究的衡量标准

宏观上可以形成产业并达到规模的原始创新技术毕竟是有限的，而大多数基础研究主要是对现有技术的不断改进与拓展。以一项基础研究是否形成产业来衡量它的成功与否固然简单，但这种方法却很容易抹杀其他绝大多数基础研究推动产业发展所作出的贡献。虽然如何去衡量一项基础研究是成功与失败非常困难，但大多数情况下，我们至少还是两种常见的标准用于衡量：一是论文引用数；二是对产业贡献。

我们以计算机体系结构领域来探讨这两个标准之间的关系。体系结构领域的基础研究成果主要发表在学术会议 ISCA (International Symposia on Computer Architecture) 上。1998 年在 ISCA 创办 25 周年之际从 1973~1997 年的 900 篇 ISCA 论文中挑选出 41 篇精选论文⁶（占 4.6%），这些论文被认为是对计算机体系结构发展有重大贡献。我们对这些文章的引用数目分析⁷，发现它们的平均引用次数达到 160.8 次，是所有 ISCA 论文平均引用次数 20.3 的 8 倍，这些论文的引用次数在所有 35 年 ISCA 论文中更是占到了 23.7%。但是通过对这 41 篇论文的单篇引用次数（图 5）分析，我们会发现论文的引用次数分布极度不均匀。引用最高的《Active Messages: A Mechanism for Integrated Communication and Computation》达到了 893 次，而最低的《Computer Structures: What Have We Learned from the PDP-11》仅有 5 次。这是为什么呢？从 John Hennessy 的报告《ISCA 25: Looking Backward, Looking Forward》也许

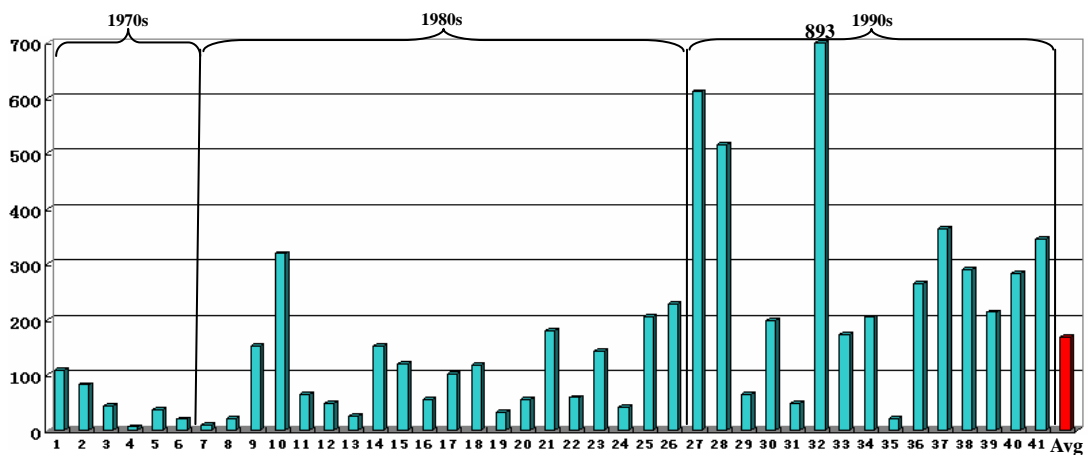


图 5. 41 篇 25 年 ISCA(1973~1998)精选论文的引用数目分布

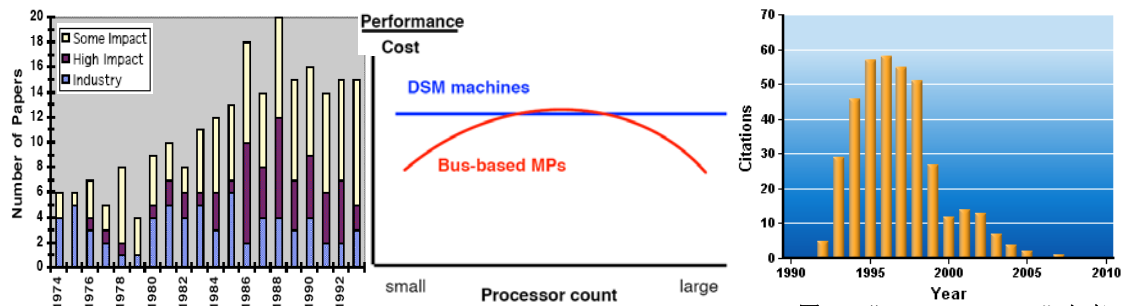


图 6. ISCA 文章的影响力

图 7. 分布式共享内存(DSM)系统的兴起源于性价比

图 8. “Active Message”文章引用次数历年分布图

⁶ 25 Years of the International Symposia on Computer Architecture, ACM,1998.

⁷ 论文引用次数数据来源：<http://libra.msra.cn/>

可以找到一些原因。图 6 是该报告中对 ISCA 文章影响力的分类与统计。他把 ISCA 论文按照产业关联程度 (Industrial Relevance) 分为 3 类: 源自产业, 高影响力和一定影响力。应该说, 上述 41 篇精选论文都属于源自产业或高影响力的论文。我们将 41 篇文章按照来自大学还是企业进行分类 (见下表), 发现大约有 29% 的文章与企业相关, 它们的平均引用次数约 143.7 次, 若剔除两个特例 (最高与最低), 平均引用只有 120.3 次, 分别低于大学的 179.2 次与 158.6 次。这一定程度上说明企业研究更注重短期目标, 其论文引用数低一些。大学依然是基础研究的主力军 (占 71%), 相对来说更注重长期性, 因而其论文引用数要高。但是当大学与企业合作, 服务于企业短期目标时, 其论文引用次数则大幅下降。因此, 成功的基础研究应该能引领企业技术的发展, 而不是被企业短期目标牵着走。

	文章数	占百分比	平均引用数	剔除最高最低后平均
企业	7	17.0 %	154.1	111.6
大学企业合作	5	12.2 %	129	120
企业相关	12	29.2 %	143.7	120.3
大学	29	70.8 %	179.2	158.6

另一个有趣的现象的是七八十年代论文的引用数目远小于九十年代, 一方面原因是大多数企业研究 (有 9 篇论文) 集中在那段时间, 我认为另一个重要的原因是那时正处于微体系结构的萌芽期, 从发明新技术到企业应用该技术周期很短, 并没有经过反复研究的观望期。然而随着微体系结构进入了成熟期, 能直接被企业应用的原始创新越来越少。不过随着研究人员数目的增加, 且多机系统处于萌芽期, 90 年代又出现了新的技术创新潮。

讨论一项基础研究是否成功不能脱离它所处的时代背景。比如分布式共享内存系统 (DSM) 的兴起源于它诱人的性价比, John Hennessy 的报告中 (图 7) 描述 DSM 与基于总线的多机系统相比的性价比优势。但随着 “Memory Wall” 问题的加剧, 本地访存都已经成为系统性能的瓶颈, DSM 的性价比优势也随之被打破, 那么它也便失去了市场。但我觉得, 我们并不能因此而否认 DSM 的成功 (当然, 从事基础研究的人们肯定不希望自己的技术只是昙花一现)。当一篇有影响力的论文发表后, 它的思想往往很快就能被其他研究人员所捕获并发展, 图 8 是《Active Messages: A Mechanism for Integrated Communication and Computation》在 1992 年发表后历年引用情况, 很好的反映了这一点。由于体系结构领域技术更新非常快, 一些技术从发明到成熟的周期很短, 因而很少有一篇文章会在发表 10 年后才被人认识到其重要性 (Transactional Memory 等是例外)。

研究成果的影响力与研究方向固然相关, 但判断研究成功与否则与研究方向无关。对此, 微软副总裁 Rick Rashid 也曾发表过类似见解: “你不能跟研究员说他们该做什么, 这种方式不可能产出好成果。你只能尽量去创造一个环境, 去雇用最好的人, 给他们机会去将研究能力进一步提升。微软不能决定研究员做什么, 只能研究员决定自己要做什么。从历史上来看, 惟其如此, 假以时日, 研究者才可能有优秀的成果。微软可以开除某个研究员, 但只会是因为他的研究工作很糟糕, 而不是因为他的研究方向错了。”

总结来说, 由于计算机学科与产业紧密结合, 且优秀的 IT 企业一般都具有企业自身研发能力与跟踪大学基础研究成果的能力, 因此, 对产业有推进作用的基础研究往往也是论文引用数目较高的研究。这也便是我认为上述的两个指标 (论文引用数目与产业贡献) 存在的内在联系: 一项基础研究成果的论文引用数目可以反映出该研究成功与否, 且在一段时期必然起到推动产业的作用。

对我们基础研究的启发

持什么观点?

对于基础研究，我们该持什么观点？现在主要有两种观点，

- 一种是**学术领先并不会对社会、经济产生影响**；
- 一种是**学术领先才能达到产业领先**。

个人倾向于第二种观点，而且我们在讨论上述观点时更不能忽视一个前提，即在“学术领先”的基础上。因而，首要的是如何达到“学术领先”？对于这一点，一些观点认为由于我们人均科研资源有限，中国的基础研究更应该重视有效产出，而不能简单地学习国外的“Paper-To-Paper”模式。

美国政府拥有充足的科研资源，允许科研人员从事各种研究，但却未必是最有效的资源利用方式。虽然在报告《Innovation in Information Technology》中列出 19 个成功的基础研究到产业化例子，但我们也可以举出更多无效研究（“Paper-To-Paper”式等）。之前已经分析过计算机学科是“人造学科”，非常容易创造新方法，但也更容易浪费科研资源。如果仅仅把科研目标局限于复杂设计空间中的“某种特定条件、指定子空间”，就很容易产出毫无实际意义的新方法。

创造什么环境？

我们的基础研究要为国际同行所关注，那么建立良好的学术环境也是相当的重要。以下几个方面应该是良好学术环境的部分体现：

- **学术影响力评价**：对于这个问题，David Patterson 等人在《Evaluating computer scientists and engineers for promotion and tenure》有相关评论。Patterson 等认为，发表文章只是一种标准，而同行（包含重要企业）之间的“评估信（letters of evaluation）”也是很重要的。
- **鼓励合作**：国外大学教授之间合作很多，但我们国内的合作还不深入。计算所里也已经开始有一些规范鼓励课题组之间合作⁸，不过还需要其他方面配套措施，比如“跨课题跨实验室合作的主要障碍不是论文贡献而是课题和经费的管理方法”，“学生毕业要求第一作者的论文、专利可能会影响博士的毕业”⁹等。

- **学术会议发表**：计算机基础研究人员普遍认为会议比期刊更重要，因此国际同行相对更关注计算机领域会议。图 9 是由美国科学院发表的报告《Academic Careers for Experimental Computer Scientists and Engineers》对部分计算机会议与期刊的对比。从图中可以看出计算机会议从录用到发表的周期普遍要比期刊短，而且引用的文献也比期刊要更新。对于有声誉的会议或者期刊，发表周期越短，越能

	Time to Acceptance (months)	Time to Publication (months)
Conference^a		
PLDI	2.3	7.3
ISCA	2.6	6.2
ASPLOS	2.5	7.0
SOSP	3.0	7.8
OOPSLA	2.6	7.2
SIGGRAPH	2.3	6.8
Average	2.5	7.0
Journal^b		
TOPLAS	24.0	32.4
TOCS	17.0	21.3
IEEE TOC	N/A	32.2
IEEE TSE	N/A	29.2

Publication ^a	Fraction of References to Papers Less Than Two Years Old (%)	Median Age of Reference (years)
ASPLOS (Conference)	32	3
ISCA (Conference)	35	3
PLDI (Conference)	26	4
TOCS (Journal)	20	4.5
TOPLAS (Journal)	12	5

图 9. 会议与期刊的对比。(a) 发表时间对比；(b) 参考文献年限

体现该期刊会议把握研究最前沿的能力。如物理界的《Physical Review Letter》对于

⁸ 《所务会纪要-08-06》：“……对联名发表的学术论文作者，每位作者可按一定的权重比例计算贡献……”。

⁹ 以上观点主要来自 pae 平台：<http://pae.ict.ac.cn/>。“所务信息”频道，《所务会纪要 08-06》主题。

最前沿的研究成果投稿，甚至只需 2 星期即可发表。所以，我们的基础研究工作要得到同行关注，在有声誉的国际会议发表与交流是重要的途径之一。

- **学术交流：**学术交流可从内部与外部两方面改善。于内部而言，类似与国外的 Seminar 交流模式值得借鉴，并且要扩展到课题小组之间。通过一部分资金提供交流场所、资料、食物等可以吸引更多人参加。现有的学术交流平台可变被动为主动模式，像 mail-list 那样有更新便主动向感兴趣的人发送。所级学术报告除了统一通知外，可考虑提供录像、ppt 等资料供没有参加的人员了解，并且可和学术交流平台整合，鼓励大家报告结束后对报告内容进行讨论。于外部而言，除了邀请国内外知名学者来所交流，也可多与周围院校联系，共享学术交流信息。
- **博士研究生的培养：**博士研究生具体创新的主力军，因此尤其要重视博士研究生的培养。但，一方面由于博士入学门槛相对较低，另一方面由于日后的就业压力，两方面结合导致很博士并不能安心参与基础研究，不容易在就读期间发表创新成果，博士培养效率并不高。因此，改善博士培养也许可以从上述两方面着手：如适当增加直博生的比例，加强对博士毕业生就业指导（提供专家署名推荐信等），利用毕业学生校友会这份宝贵资源为应届毕业生提供更好的就业机会。另外，博士研究生作为人才梯队（研究员→博士研究生→硕士研究生）中最重要的一环，尤其要重视培养博士研究生的领导能力。可考虑建立研究员领导若干名博士研究生、博士研究生带领若干名硕士研究生的管理模式。

当然上述的几点只是良好学术环境的一部分，要建立完善的学术环境还有更多的方面需要考虑。

总结

本文通过对计算机发展历史上的 19 个关键技术进行分析，总结了计算机基础研究与产业化的关系，即：（1）产业化离不开基础研究（大学基础研究与企业 R&D）；（2）大学基础研究与企业 R&D 互动反馈周期约 3~5 年；（3）从基础研究到第一代产品大约需要 5~10 年；（4）从基础研究到 10 亿美元规模的产业化大约需要 10~15 年。并进一步分析计算机基础研究的周期长、实验工程型、非预期结果、技术交叉发展、易产业化等特点以及基础研究的两个衡量标准（论文引用数目与产业贡献）之间的关系。在分析的基础上，阐述了对我们开展基础研究的启示。