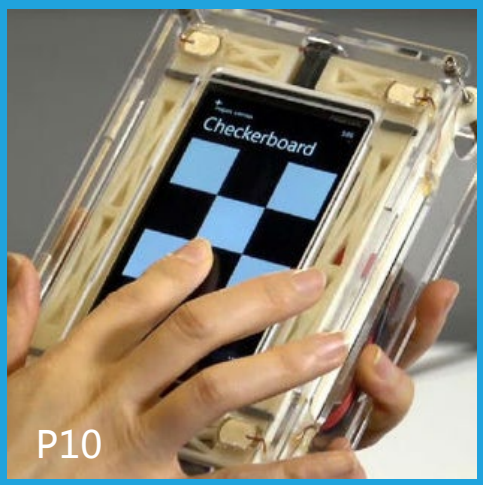
Matrix

2014年7月 第31期

P4

Microsoft Research



P10

Skype Translator: 梦想照进现实

P16

P31

Cortana : 精彩待续

P24

P34

Matrix

微软亚洲研究院 2014年7月 第31期

院长寄语 Foreword

为青年学子搭起一个舞台 3

研究院活动 Culture

编程之“美”：连接知识与现实的桥梁 4
南北编程小将携手挑战现实难题 7
共筑梦想：中美青年创客大赛 9

媒体报道 In the Media

Haptics技术：在触屏上实现物理按钮效果 10
对话电池技术专家Ranveer Chandra：不要等待革命性突破 12

微博拾粹 Weibo Highlights

@微软亚洲研究院 官方微博精彩选摘 14

前沿技术 Technology

Skype Translator：梦想照进现实 16
IE11阅读视图：带给你静心饕餮阅读大餐的片刻 19
你说英语够地道吗？ 22
Cortana：精彩待续 24
破译手势在对话中的意义 26

研究院博客 MSRA Blog

科研不是比赛，而是一种对未知和完美的自我追求 28
漫谈2014人机交互（CHI）大会 31

为青年学子搭起一个舞台

在微软亚洲研究院，活跃着这样一群人，他们朝气蓬勃，多才多艺，充满理想，崇尚Work Hard, Play Harder——对，他们就是微软亚洲研究院的实习生。正是与学生们的朝夕相处，让我们一直在思考，如何能为更多有志于计算机科学的学生搭起一个舞台，来鼓励他们学习、精进、发挥创意；既能够助力他们个人的成长，在未来也能贡献于学科和社会的进步。

第三届“微软编程之美挑战赛”在5月份刚刚落下帷幕。今年大赛的主题是“城市计算”（Urban Informatics），我们希望参赛学生能利用真实的城市数据，去尝试解决现实中存在的问题。不同于单纯的算法、编程技巧的比拼，学生们需要排除实际问题包含的干扰因素，从工程的角度不断优化自己的作品。我相信这样一种真实的研发体验，能让大家更真切地感受到编程的价值，领略编程之“美”。

今年“编程之美”大赛的冠军是来自中国一南一北的两位小将。比赛都有胜负，但我们更希望“编程之美”大赛在过程中能够真正促进学生之间、学生与企业、研究机构之间的交流学习。大家能找到志同道合的良师和益友，切磋技巧、互相启发。

5月至7月，作为第五轮中美人文交流高层磋商机制配套活动之一的“中美青年创客大赛”在中美两国多地举行。微软与同济大学共同承办了中美创客大赛上海站的选拔赛。微软为参赛选手提供了Azure和Kinect的设备支持，微软的工程师团队还为他们做了用户体验、云计算等主题的培训，帮助学生们将创意变成真正的作品。

8月中旬，我们位于中关村丹棱街5号的大楼里又将迎来一批学生参加一年一度的微软学生夏令营。每一年的夏令营活动都会出现让我眼前一亮的学生作品。今年我们会带领他们重新认识城市、认识科研，也期望学生们仍旧能用他们的创造力感染我们。

微软亚洲研究院院长

洪小文

编程之“美”：连接知识与现实的桥梁

2014年5月21日，由微软主办、电气电子工程师学会（IEEE）协办的第三届“微软编程之美挑战赛”落下帷幕，来自清华大学的王杉与复旦大学的孙祥彦获得本届冠军。之前两人素不相识，决赛当天在短短三分钟交谈后选择了彼此，临时组队，遭遇各种不顺，不过最后还是因程序运行结果最为准确而夺冠。比赛期间的惊心动魄让他们至今都觉得回味无穷。

在比赛中通过编程将想法变为现实，运用逻辑和抽象思维解决问题是他们最有成就感的地方，虽然比赛时间相当紧张，但他们都很享受整个过程。在本次大赛针对1000名复赛选手的答卷中，超过60%的选手表示实用性和逻辑性是大赛最吸引他们的地方。

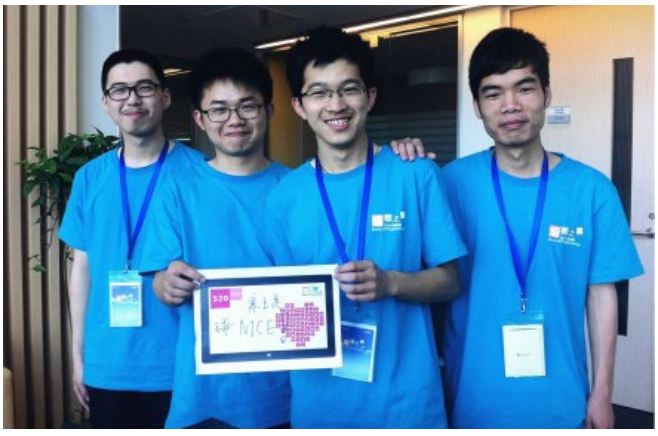
利用真实数据、解决现实问题是本次“编程之美”大赛与其他同类比赛最大的不同，让大学生在提升编程技巧的同时，更能在解决现实问题的过程中享受编程的“美”。这不单是一场比赛，更是一个交流和学习的平台，不仅为来自天南海北的学生之间提供交流的机会，也在学生与企业之间架起桥梁，为学生创造了与一线研发专家亲密合作以及更多与业界交流的机会。

享受编程 之“美”

提起编程，很多人会联想到一行行密密麻麻的代码，“美”之于编程似乎无从谈起，但对于享受到其中乐趣的人来说，编程却是充满美感的。《编程之美》一书的作者，微软首席开发经理邹欣对此颇有感触，在他看来，可以从四个方面去理解编程之“美”。

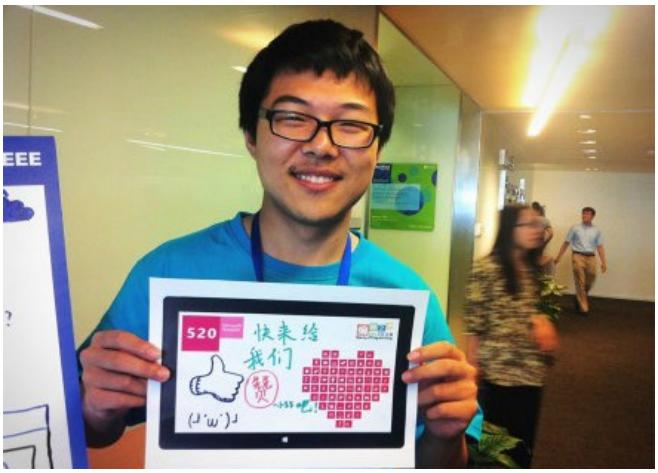
首先，程序本身一定要“美”，增之一分则太长、减之一分则太短，如美好的女子一样要恰到好处，若有冗余会影响程序运行效率，若有缺失则无法解决问题；其次，在解决实际问题的过程中展现美，当用精妙的算法去解决实实在在的问题，所带来的成就与喜悦自是难以言表；再次，编程之“美”体现在平衡，在有限的时间内做出最接近用户需求的程序，同时确保体验最佳，期间算法与代码的取舍在于平衡；最后，美是一个磨砺的过程，无论多么伟大的产品和项目，工程师们都需要经历默默无闻、潜心钻研的过程。直到有一天，就像泰戈尔诗句里所说的那样：水的载歌载舞，使他们的“鹅卵石”臻于完美。

如何在看似乏味的编程过程中做到载歌载舞？邹欣强调说，一定记得编程学习是一个自然的过程，不要人为斩断，不是老师规定考什么才学什么，时刻想着自己是要用程序去解决问题，让问题带着自己去学习，才能乐享其中。与大学环境不同，在微软这样的企业，所有的编程



都与实际问题相结合，同事之间相互鼓励、共同磨砺的协作氛围浓重，微软也是希望借助大赛让更多大学生体会到编程之“美”。

在邹欣看来，非计算机专业的学生其实也同样能够享受到编程之“美”，编程当中所体现的计算性思维（Computational Thinking），不仅是从事计算机科学的人需要的，也是年轻一代每个人都需要的。微软全球资深副总裁、微软研究院的周以真（Jeannette Wing）博士曾说，计算性思维是人类的一项基本技能，世界上每个人都在使用，我们甚至应该把它纳入教育计划，让它像阅读、写作、算术等一样，成为每个孩子都应具备的基本能力。



可以认为，编程和计算性思维更关乎一种解决问题的思维方式，年轻人完全可以基于兴趣从小培养。对于已经充分享受到编程之“美”的上届冠军李雪来说，编程是一件很自由的事情，像写作，每个人都有自己的表达方式和对于问题的理解，每个人解决问题的过程都截然不同，基于对问题的不同理解，充分发挥创造力，表达出不同的结果，正是编程之“美”所在。

解决现实问题，模拟真实环境

此次大赛的主题是“城市计算”，在决赛中采用了由江苏信息服务产业基地提供的来自扬州市的真实数据以及环保部所公布的气象数据。选手们可以基于这些数据建模，并从工程的角度去不断优化，选手们不仅要有强大的逻辑思维能力，还要兼顾实际问题中的各种干扰因素，最终导致解决问题的方法迥异，却也最能体现选手的创新能力。

孙祥彦此前曾多次参加ACM大赛，在他看来，ACM大赛中都是算法的较量，在学校的训练也多以算法题为主，大家对城市计算并不了解，因此参赛之前，他特意查阅了城市计算相关领域的研究，其中包括微软亚洲研究院的实时动态拼车系统和实时空气污染监测，这才深刻体会到算法理论与工程项目之间的差别有多大。应该说，这是他第一次真正利用编程来解决实际问题，比赛过程很艰难，却让他收获颇丰。他和王彬非常期待之后能到微软实习，真正进入“实战”的环境。

“我们为编程之美大赛提供扬州市的第一手数据，是希望让参赛

的大学生们身临其境，感受到城市发展面临的真实环境，从而激发他们的创造力，真正尝试去解决实际问题。”江苏信息服务产业基地管理办公室主任陈栋表示：“同时也希望通过这种类型的比赛，探索利用大数据、云计算等新兴技术为城市发展和规划找到一条行之有效的可持续发展之道。”

大赛不仅决赛主题与现实问题挂钩，其形式也模拟真实环境中的项目创作，很像是创业企业成长的缩减版，充满挑战。56名决赛选手在一个房间里随意交流，通过十组五分钟交流选择自己的合作者，在随后两天的时间里一起开发、一起面对评委演示和答辩，就像是创业企业面对VC一样，完全是高手之间的协作与较量，选手们的沟通能力和团队协作能力在每个环节都展露无遗。

与就业接轨，重在经验积累

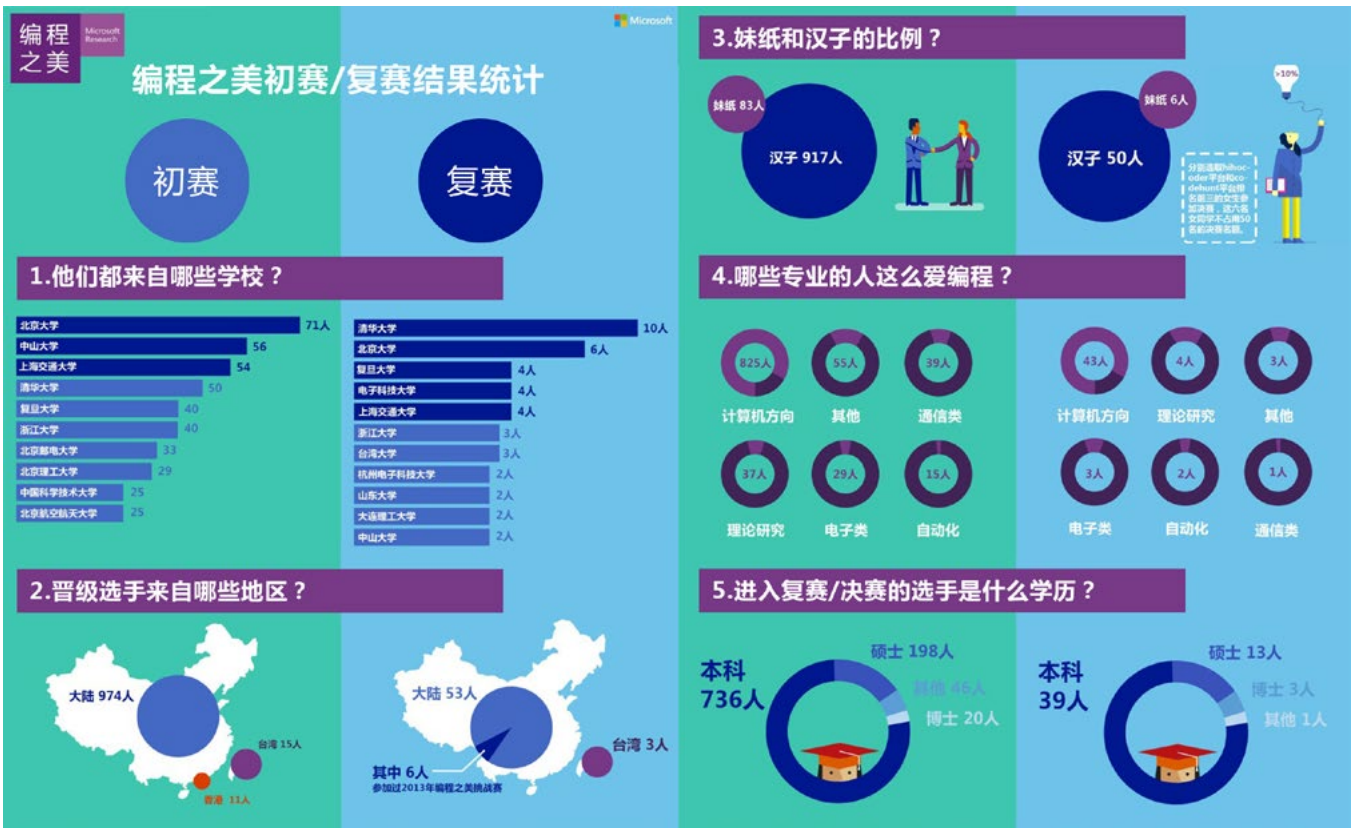
与一线专家面对面的交流，体验微软文化，乃至日后进入微软实习，是决赛选手的另一大收获。大赛的评委既有来自微软的研究员和工程师，也有来自高校教师和扬州市的政府代表，选手们可以从他们这里得到最中肯的意见和建议。

自2012年举办第一届编程之美大赛以来，已有20余名优秀参赛者获得了到微软实习的机会，在研究员和工程师的指导下参与前沿的科研研究和实际的工程项目。王彬说，他当初决定参赛的唯一原因就是能有机会来微软实习，中学时期开始对编程感兴趣的他一直使用微软平台，微软工程师改变世界的梦想也是他的梦想。

与学生就业难相对应，企业寻找人才也同样难，之间的鸿沟需要更好的解决办法，大赛因此成为连接知识与现实的桥梁。作为一种另类的“实习”经历，大赛让学生在校期间就有机会与国际领先企业面对面。哪怕没能进入到最后的决赛，过程本身便已是经验和实践的积累与丰富。正如上届冠军李雪所说的，对于学生而言，“编程之美”不单纯是一场比赛，更是一个与更多热爱编程的伙伴交流和学习的平台。

IEEE的统计数据显示，中国的IEEE学生会员现有三四千人，而在印度这个数字是中国的十倍，从这个侧面也可以看出，中国大学生与IT业界的联系不够紧密。在IEEE国际业务发展总监华宁看来，参赛学生通过“编程之美”认识到编程的重要性，提高了自己编程的技巧，进入社会后为国家科技的发展和自主创新做出更大的贡献，这正是IEEE作为一个非盈利机构非常愿意与微软在这个项目上进行长期合作的原因。此外，IEEE还将提供更多中文本地化内容，帮助中国的大学生克服语言障碍，与更多业界的国际精英实现交流和沟通。

“在微软看来，像‘编程之美’这样由企业组织的比赛越多越好。比赛的输赢不重要，了解最新的技术与方法论，了解新兴技术与现实问题的结合点，对大学生来说才最有帮助。”微软亚洲研究院学术合作部总监潘天佑博士表示：“企业需要能够解决实际问题的人才，此类大赛恰恰能够为学生们创造一个实战平台，让他们提前接触业界，学以致用，为将来的就业和发展做准备。”



南北编程小将携手挑战现实难题
——记2014编程之美挑战赛冠军孙祥彦和王杉

编者按：2014编程之美挑战赛冠军是来自一南一北的小将：孙祥彦(复旦大学)、王杉(清华大学)。在初赛和复赛，两人巧合选择了寓教于乐、注重逆向算法的Code Hunt平台；决赛中，二人组队攻下预测空气质量的难题。这背后还有什么有趣故事呢？

5月21日，微软2014编程之美挑战赛完美落幕。来自复旦大学和清华大学的本科生孙祥彦和王杉在历经多轮挑战后，以其出色的编程能力和运行结果的准确度问鼎冠军。

一南一北的两位小将在地域上看似没有任何交集，但却在编程这件事上默契十足。孙祥彦今年大四，热爱代码，从小学四、五年级就开始编程，聊起编程，总能侃侃而谈；王杉今年大三，比较内敛，不像常见的北京人那么爱侃，但在代码、算法面前他却能偶尔收起羞涩。

独僻蹊径，“游戏”过关

今年的编程之美比赛有两个比赛平台：一个与往年的比赛平台类似，侧重算法；另一个却是今年别出心裁的尝试，引入寓赛于乐、侧重考察选手们的逆向算法实践能力的Code Hunt平台。出于不走寻常路、挑战自我的考量，在赛事初期，孙祥彦和王杉不约而同地选择了Code Hunt，像玩游戏一样有趣、好玩是他们对该平台最深的印象。

Code Hunt是基于Microsoft Azure、由微软研究院开发的在线编程与评测系统。多元化的测试与评价体系既为来自不同背景、拥有不同技术偏好与擅长领域的参赛者提供充分展现自身能力与特长的平台，也希望能够培养参赛者双向思维的能力。

“我是被比赛的多样化吸引来参赛的。在初赛和复赛的时候，它有两个不一样的平台，我出于好奇就选择了Code Hunt。现在回想一下，这个平台其实特别有趣。”孙彦祥笑谈道。

在谈到自己过关斩将的秘诀时，孙祥彦总结了一个小贴士分享给大家。“比赛之前，要好好了解比赛规则”。他说道：“学计算机的人，很多时候不喜欢看很长的文档，但就是这个习惯，很可能会令你错失一些机会。比如，在Code Hunt平台的比赛规则中，有一条计分规则是每道题第一个做完的人会有加分。复赛还是比较紧张的，大家分数都很接近，一不小心可能就落下了。当时我就打乱顺序做题，加了一点分。”王杉也同样注意到了这一点。

的确，只有了解游戏规则，才能更好地玩游戏。

预测空气质量，挑战频现

5月20号，50多名经过多轮选拔的决赛选手抵京展开终极较量。大赛组委会采取了两人组队的决赛形式，并提供了两道题目让选手选择——一道是利用扬州的真实数据真实场景，鼓励倡导参赛选手开发针对民生方面的应用与服务，不拘泥于PC、手机、平板等终端访问方式，凸显任何时间、任何地点、任何人享受智慧城市的信息化应用与服务，另一题则是以时下大热的“城市计算”为主题，采用环保部公布的真实数据，要求选手们从工程角度出发，预测不同区域的空气质量。

怀着触摸现实挑战的想法，两人再次不约而同地瞄准了“城市计算”这道题，然而比赛过程却险象环生。要实现对空气质量的预测，需要兼顾到现实世界中的多种影响因素，如温度，气压，风速等。在建模的过程中，一个因素考虑的不周详都有可能导致最终的结果不尽如人意。

“已经到了下午三点了，还没有跑出结果，内心特别着急。”孙祥彦回忆道，“我想用神经网络来解答这个问题，但是因为工具不是特别熟悉，调试到下午3点多，代码还是完全没办法工作。”

决赛持续12小时，从清晨的早10点到月上梢头的晚10点。而下午三点意味着比赛时间已经流逝了差不多一半。他们回想道：“当时我们就想，不管怎么样，我们至少要跑出一个结果。”

最终得到了还算满意的结果，但不料在第二天的展示环节，所采用的方法却又屡屡被在场的评委提问质疑。最后，经过与评委的反复沟通，终于获得肯定。

两人在算法和编码方面的各执一长，也对他们俩的夺冠起到了重要作用。“其实我们两个还是很合适的，王杉相对对算法比较了解，而我的编码调试能力更强一些，很互补。他可以给予很多算法上的帮助，而我可以很快把它实现出来。”孙祥彦补充道。

本次比赛的评委郑宇博士这样评价他们最后的作品——“这两个同学提出的方法考虑了气象条件和周边空气质量监测站点的各种信息，并利用SVM来做一个循序渐进的推断。这个方法实际上是同时考虑了空气质量在时序和空间相上的变化，因此取得了比其他同学更高的预测精度。方法虽然并不复杂，但思维清晰、敏锐，准确捕捉到了问题的关键，小有庖丁解牛的韵味。”

郑宇还谈了谈本次大赛选择空气质量作为题目的三大缘由，他说道：“首先，在以往的各种程序设计大赛中，同学们遇到的题目通常都预先存在一个唯一的最优解，题目留给同学们设计自己方法的空间较小。本次大赛希望区别于其他的程序设计大赛，考察同学们解决不确定性问题的能力。空气质量预测是个很新颖的课题，暂时没有成熟的解法，并且不存在一个唯一的最优解法。这样会留给学们更多的思考空间，激发他们的创作力和想象潜能。其次，在大数据时代，本次大赛也想考察同学们对多种数据源的分析和处理能力。让他们小试牛刀，感受大数据的奥妙。最后，空气质量是当前大家关注的热点，关系到民生。这个题目具备一定的公益性，能加强大家对空气质量现状的了解，并唤起同学们及更广泛群体对环保的关注。”

不同编程故事，一样的陶醉与热爱

孙祥彦今年大四，却与编程有着颇深的渊源。

小学一年级的时候，爸爸给家里买了一台电脑，孙彦祥开始接触玩电脑。到了小学四、五年级，有老师带他去参加了小学信息学奥赛。虽然这只是县里面的一般比赛，当时用的只是计算机Basic语言，但这让编程的种子在他内心萌芽。后来，从初中到高中，即使学业再忙碌，他都没有放弃过编程这项爱好，一直坚持参加信息学竞赛。他说：“高中学习压力大，但每天晚上还是会去训练2、3个小时。直到后来到了复旦，加入ACM队，泡在编程上的时间就更多了。”

除了爱参加程序竞赛，一有闲暇时间，孙祥彦还特别喜欢去做一些开源项目。似乎对他而言，做开源项目已是极大的娱乐——“看着一行一行的枯燥无趣的代码，换成了那些很有趣的应用或产品，内心特有成就感。”

而来自清华大学的王杉跟孙祥彦的编程之路走得不太一样。

在2014编程之美挑战赛前，王杉专注在学校专业的计算机知识修习上，没有参加过任何编程比赛，微软的平台成为了他的第一个试金石。他谈道：“我一直都向往能够有机会来微软实习，今年一看到编程之美的比赛信息就毫不犹豫地报名了。”虽然此前没有参加过编程比赛，但在学习中所掌握的扎实的代码基础，让王杉在这次比赛中一路畅行。

成长的故事炯然不同，但对编程的喜爱却十分相似。一场比赛让两人与万名参赛者聚在一起并肩作战，并脱颖而出。孙祥彦视这为一段美妙的经历，王杉也十分激动自己进入微软实习的愿望终于要实现了。大赛圆满落幕，但关于编程的故事不会止息。下一年，又会有怎样的故事上演呢？敬请期待。

共筑梦想
——中美青年创客大赛

“创客”来源于英文“Maker”，指努力把各种创意转变为现实的人。由中华人民共和国教育部国际司主办、教育部留学服务中心、清华大学及新奥集团承办的“共筑梦想——中美青年创客大赛”是第五轮中美人文交流高层磋商的系列活动之一，旨在激发中美两国青年的创造力和创新力，为两国青年搭建了互相交流、互相学习和互相促进的平台。

该大赛自5月6日在京启动以来，先后举办了北京、上海、深圳及成都选拔赛。此次中美青年创客大赛上海站于5月底举行，由同济大学与微软共同承办，比赛吸引了150余位来自中美两国极富创客精神的青年，大家齐聚同济大学设计创意学院，进行了4天创客马拉松之旅。参赛选手分为30组，创意想法聚焦于应对环境变化、改善人类生活及促进社会发展的各个层面，并运用各种最新的技术手段，经历72小时的脑力激荡和高强度的开发过程，在短短数天内将创意化为现实。

在比赛过程中，微软工程师组成了强大的导师队伍，为选手们带来了关于用户体验研究、云计算、Kinect人机交互设计的培训，同时指出了研究与设计时存在的错误与认识误区。以微软的商用公有云服务为实例，展示了不同的研究与设计方法在实际微软产品研发过程中的应用。Microsoft Azure不仅是协助企业更专注于核心商业运营的利器，也是让学生及学校发挥创作的最佳实践工具。

导师还为选手们详细讲解了Kinect在人体识别、3D扫描、实时建模方面的强大功能。现场创客热情参与了台上互动进行体验，多个队伍在导师指导下选择Kinect作为他们实现自然用户交互的平台。除此之外，微软还为选手们提供了免费云平台账号，让选手更加自由、流畅地在云端进行大数据的整合及使用。在同济大学设计创意学院创意工场，选手们自如地挥洒灵感，将一个个大胆的创意变成眼前的设计作品。各组注重用户体验、设计、拥有明确目标的设计原型正在逐渐成熟。很多团队的产品概念都拥有鲜明社会价值。

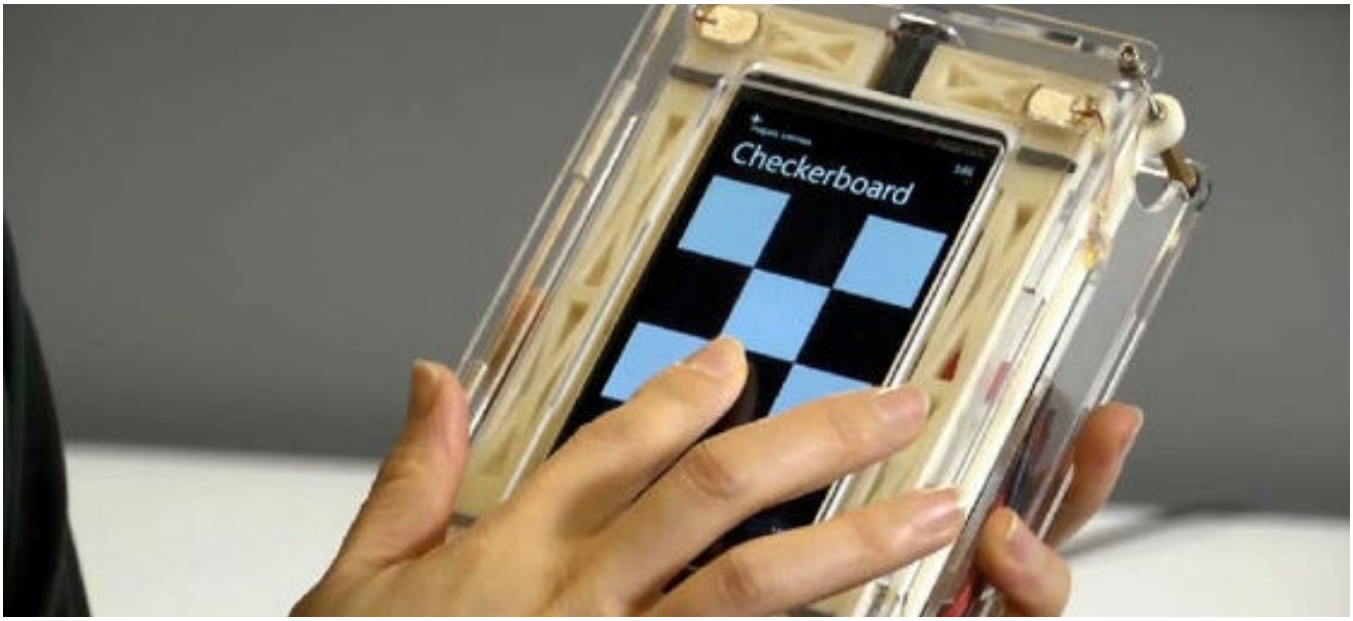
7月初，中美创客大赛决赛在河北廊坊新奥集团举办。100多名中美青年学生组成30支队伍进行创客比拼，经过了48小时创客马拉松角逐及来自不同领域评委的认真评审，最终评审出两组一等奖、三组二等奖及五组三等奖。10日上午，作为第五轮中美人文交流高层磋商机制配套活动之一，中美青年创客大赛颁奖大会暨大学生创新创业论坛在人民大会堂重庆厅隆重举行。本次大赛的主办机构、承办机构、中美青年创客代表及企业代表160余人参加了此次大会。郝平副部长和美国施坦格尔副国务卿参加了颁奖大会，共同观看了大赛视频，为获得本次大赛一等奖的团队颁发了获奖证书与奖杯，并分别做了精彩讲话，高度评价了两国青年在本次比赛中的优异表现，充分肯定了此次活动对中美青年人文交流的重要意义。

上海赛区的漂流瓶项目获得中美青年创客大赛决赛一等奖，该项目利用现有的开源硬件平台，制造出服役期10年、低成本的新型漂流瓶。预期功能是定位自身位置、收集周围洋流及物理信息、现实通讯等；主要用途是辅助气象部门分析、海洋研究；未来可能发展方向是传感器校准、地面和水下视频、波谱、海啸预警等。基于这一平台，也有开发出多种数据及商业应用的前景。施坦格尔副国务卿特别询问了微软Azure云平台在项目中的应用。Azure提供业界顶尖的云计算技术，能够与企业现有本地IT设施混合使用，为选手能在决赛期间进行数据的收集、分析提供强大支持。

中美青年创客大赛不仅有创意的碰撞，也是个性的展现。不同国籍、地域与身份的创客们汇集一堂，他们享受着融洽欢快的团队合作，而差异碰撞出的有趣火花也同样引人关注。我们也十分期待看到这群充满活力的青年创客们能带着自己的创意与热情飞得更高，走向更加广阔的舞台。

Haptics技术：在触屏上实现物理按钮效果

《新浪创事记》　文/张虹



“我觉得自己漂浮在床上，却又感受不到床的存在，我发现偶尔自己坐了起来，但马上又倒下了——因为我感受不到身体。”

19岁那年，一种罕见的神经系统疾病夺去了Ian Waterman脖颈以下的全部触觉和对肢体空间位置的感觉。Ian知道损坏的神经无法复原，因而只能绕过它们，寻找一种新方式控制自己。他穷尽所能观察、记忆、练习每个动作应当如何触发。普通人下意识就能自然完成的动作，对Ian来说，都离不开双眼观察和头脑思索如何让一系列指令精确执行。一年后，虽然如木偶般踉跄，虽然他总要低着头观察脚步的移动，但Ian终于又能站起来走路了。Ian对抗命运的故事写进了《Pride and a Daily Marathon》这本书和纪录片《The Man Who Lost His Body》中。

而与Ian有着相似不幸遭遇的其他人，终生却只能与轮椅为伴——你是否曾意识到，身处二维互联网世界中的我们，在与机器交互的时候，其实并不比失去触觉的人优越多少。

指尖交互的困境

当你每天用手指在手机、平板的玻璃屏幕点来划去时，当许多人将未来的交互畅想为用手指在这些玻璃屏幕背后显示的界面上、甚至是空中操作时，你可能已经忘却了双手是我们最灵活敏感的身体延伸。

我们的双手既能感受岩石的锋利坚硬，也能体会丝绸的飘逸轻柔，能分辨沙土的松散干燥，也能区别油水的湿润粘稠。敏感的触觉让我们能清楚知道书是否多翻过了一页，即使在黑暗环境，我们也能只凭触觉系好鞋带，扣好衬衣纽扣，摸到钥匙孔打开门锁；然而在寒冷的冬日，当双手不再自如灵活，你也许体会过纵使有双眼的帮助，仍无法做好这些事。然而当我们与“玻璃下的图片”进行交互时，其实是以视觉作为反馈的唯一方式，正如同将麻醉剂注入双手，指尖对所有界面的触觉感受都别无二致——在这样的二维加视觉式的交互中，我们浪费了造物主赋予的两平米肌肤、两百个关节与六百块肌肉。而Haptics正是一门研究如何让每个人的触觉都能用于人机交互的前沿学科。

何谓Haptics？

Haptics通过硬件与软件结合的触觉反馈机制，模拟人的真实触觉体验。由于人体感受机制复杂，对Haptics技术做清晰地分类并不容易，不过从感受输入的角度，大致可以分为对表皮，以及对肌肉中感受器刺激两类。

对普通人来说，前者其实并不陌生，手机上的“振动”就是一种表皮Haptics技术，不过许多人的误解也源于此——认为Haptics等同于“振动”。事实上这项技术远非如此简单，只是因为振动技术容易实现，而且商业化产品成熟低价。振动其实只是Haptics领域的很小一部分，在许多场景下（例如按键反馈），以振动作为触觉反馈的效果都不够好。而与这两类输入对应的技术则五花八门，跨越的技术领域特别广泛。这也要求Haptics研究者需要了解许多领域的知识。以振动反馈为例，不但要从电子工程和机械学角度知道如何最有效地利用电能、设计原件结构，还需从心理学角度了解人体对哪些频率的振动最为敏感。不过，这也是从事Haptics领域研究的乐趣之一——尤其是对兴趣广泛的研究者来说。

生活中的诸多不便都可能藉由Haptics技术的引入得到解决。比

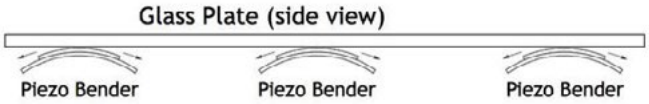
如，人们在面对面交谈时，一方看手表的动作往往会引发对方不适，倘若手表镜面应用了Haptics技术，只需手指轻轻一摸，便能知晓准确时间，这类生活中的窘境便迎刃而解。另外一个例子是网购：现实中，我们通过触摸感受材料的质地与纹理，网购则无法实现这一点，随着Haptics技术的发展，如果平板电脑将来能模拟商品的真实纹理与质地，那么网购体验将会有革命性的提升。

虚拟触觉的“RGB三原色”

正如老话说，前景总是光明的，道路总是曲折的。Haptics是一个崭新的领域，难在建立研究的基准。我们类比视觉领域的既有研究成果，尝试将纹理分解成几个独立变量。如同光的三原色一样，我们认为粗糙度、粘滞度、柔软度很可能是触觉的三个基本维度。在理解它们后，怎样用机器模拟，以不同比例混合，是否有可能根据它们定义所有的纹理呢？当前，Haptics的研究重点之一便是理解人们如何感受纹理。

粗糙度是我们研究最深，也是目前掌握最全面的技术，而粘滞度牵涉摩擦力，则相对难以实现。此外，由于三个维度往往分别单独采用不同的技术实现，而在完成最终产品时，还要考虑三者是否能够相互兼容，同时实现。在微软亚洲研究院，我们对如何在消费产品及可穿戴设备中应用Haptics技术的研究颇感兴趣，并且已开发了多种Haptics设备原型，能实现键盘的触觉反馈、在玻璃触屏上实现类似物理按钮般的效果。以键盘为例，在对照实验中，在拥有精确触摸反馈的键盘上，其输入速度和正确率都要高出单纯利用视觉反馈的键盘许多。而归类来说，这些原型中使用的Haptics技术包含以下两类。

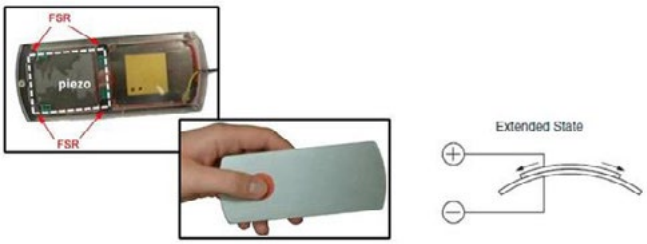
一类是利用压电效应。将压电弯曲元件置于触摸表面之下，当它们受到高压产生振动时，带动表面也同时发生上下位移。

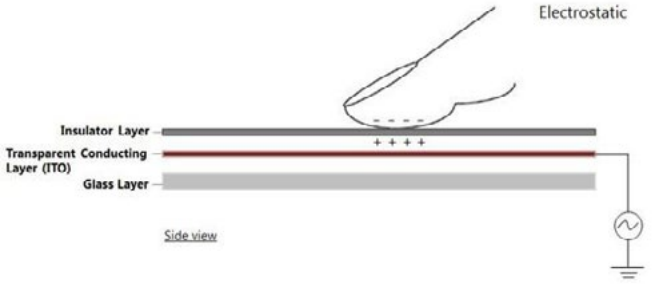


还可通过收缩表面，模拟按键咔嗒的感觉。其原理是将一块大的压电元件粘合在不锈钢片上，因为不锈钢片具有很强的面内刚度（in-plane stiffness），通过向压电元件施加高电压，元件的收缩将引起不锈钢片向一侧凹陷，在指尖产生类似轻触开关般的咔嗒体验。我们还能以超声波的频率（20-40kHz）振动触摸表面，令一层薄薄的空气附于其上，这层空气能让指尖与表面的接触部分变得特别顺滑。

另一类是静电效应。在下面的示意图中，玻璃屏幕之上还附着有两层材料。红色的是导电层（例如ITO导电玻璃），其上是绝缘层。当电信号通过导电层，手指皮肤中将产生极性相反的感应电荷，手指与屏幕的摩擦力将受到这些感应电荷相互作用力的影响。这些额外的作用力微小，我们只能体会到摩擦力发生了改变，却并不能单独感受到它们的存在。配合对手指位置的精确感知，以及相应位置电信号强度的变化，将让我们产生接触表面的粗糙程度发生了改变的错觉。

由此也可见，Haptics是一种交互，不仅与“显示”有关，还与动作捕捉及感应紧密关联。在意识到这一点之后，最近有一些技术展示，实现了极高的位置分辨率，进而能显示非常细腻的纹理，实现灯芯绒般的手感。从鲜为人知的术语，到被寄予希望的未来技术，Haptics正逐渐走向我们的日常工作和生活。Amara法则说，我们常常高估科技的短期影响力，而又低估其长期影响力。在互联网世界逐渐三维化，并越来越和现实世界水乳交融的过程里，人和机器的交互界面将不再受限于玻璃平面之下的二维世界法则，会变得跟现实世界一样细腻丰富。要让梦想更快地实现，还有赖于研究者们更多的关注和科研突破。





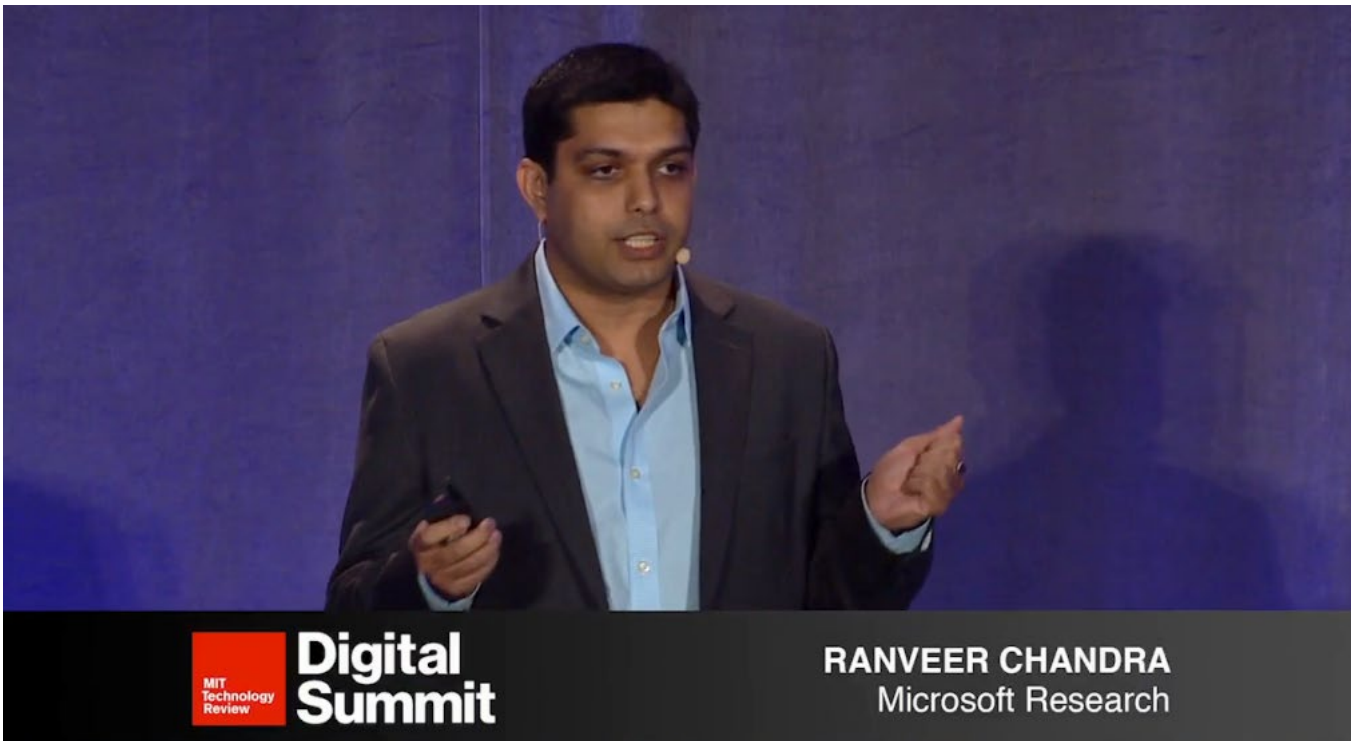
作者简介

张虹

张虹（Hong Z Tan）现任微软亚洲研究院高级研究员与人机交互组的经理。她目前也是美国普渡大学电子与计算机工程学院教授，机械工程学院客座教授，心理科学系客座教授。她在美国麻省理工学院（MIT）获得硕士和博士学位；在上海交通大学获得学士学位。她曾担任美国麻省理工学院媒体实验室的研究员（Research Scientist at MIT Media Lab）；曾获得McDonnell Visiting Fellowship在英国牛津大学访问；作为Visiting Associate Professor在美国斯坦福大学计算机系访问和工作；并曾是上海交通大学生命科学与技术研究所的客座研究员。她在基于触摸的人机界面以及触摸感知技术的研究领域享有盛誉，已经发表一百多篇具有广泛影响的学术论文，并在多个国际学术期刊编委会中担任重要职务。

对话电池技术专家Ranveer Chandra：不要等待革命性突破

《雷锋网》　文/王星



软硬件的不断进化催生大量的移动设备，不过有一项技术在过去数十年间却进展缓慢，并成为了移动设备发展的最大阻碍——它就是电池。

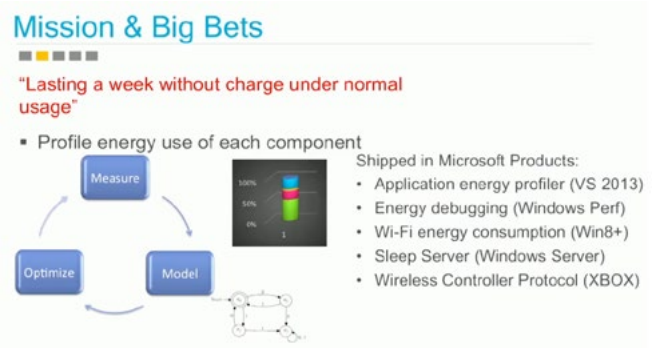
在近日于旧金山举行的《麻省理工科技评论》数字峰会上，我遇到了微软研究院高级研究员Ranveer Chandra，并围绕电池技术的发展和他进行了探讨。他在微软研究院工作了九年，专注于电池效率和无线传输方面的研究，曾被《麻省理工科技评论》评为科技创新领域的35名35岁以下青年俊杰之一。看到我在使用Macbook Air，他问我为什么用这款电脑，我提到的原因之一是它续航时间比Windows电脑更好。

“这就是我们要研究电池技术的一个原因，”他说，笔记本电脑的续航不只是系统问题，也和硬件有关：“如果你能做到垂直整合，能自己定制硬件，你就能让电池有更好的表现。”他说，电池技术是一个综合的领域，涉及到软件、硬件以及化学工程，“你在做研究时不能只关注某一个方面”。

雷锋网：为什么过去数十年电池技术没有特别大的突破？

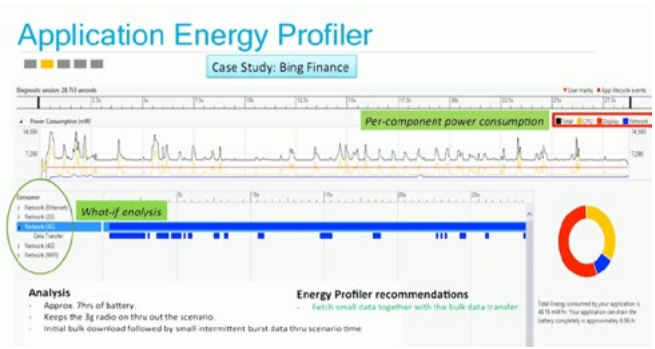
Ranveer：电池的好坏主要有两个衡量标准：一定电量电池的重量，以及一定电量电池的容量（体积）。在过去几十年里，人们都在寻找制造电池最合适的化学物质，做了很多尝试；最后锂电池出现了，它有许多好的特质。但没有哪一类电池能满足所有的需求，锂电池适合移动设备，但不适合汽车。

在那以后，人们又开始研究很多新的材料，事实上已经取得了一些突破，这主要是来自于对不同电池的组合；尤其是针对电量之外的其他特性方面，比如耐用性，如今有的电池可以充电数千次。不管怎么说，已经有很多创新正在发生，比如人们正在研究锂离子聚合物电池、固态电池；当然我们不会为此自豪，因为还有很多困难亟待解决。



雷锋网：在IT领域存在摩尔定律，电池技术方面是否有类似的定律呢？

Ranveer：很遗憾，没有。我也希望有这样的定律。目前还没有办法取得那样快速的突破，但人们正在尝试。



雷锋网：在软件层面，如何对操作系统进行优化才能在现有的条件下提升设备的续航时间？

Ranveer：我们正在做很多工作。其中之一是对进程的智能安排（intelligent schedule）。我们从微软的设备上收集了很多数据来做这方面的研究，我们希望能在后台确保可预测的电量消耗，比如你不应该在关闭屏幕后依然看到显著的电量消耗，而事实上在许多设备中会发生这样的事情。

如何确保这一点呢？我们收集的数据包括，包括哪些进程是活跃的，它们占用的CPU和网络资源，以及造成的电量损耗。当然，我们搜集这些数据是匿名进行的。

根据这些数据，如果我们在15或30分钟内看到了大量的电量损耗，就会利用机器学习等技术准确地鉴定出消耗电量的进程。我们在云端做这个工作，而云将通知系统，系统则会把这些进程放进一个沙箱里。

举个例子，如果我在下载一个更新，而网络连接很糟糕，设备的电量就消耗得很快；这个时候，系统监测到了这个电量损耗，就会借助机器学习去找到这个下载进程，然后把它放进沙箱；在接下来的15分钟时间内，它将只会获得一小段的时间配额；而在这个15分钟之后，将会有一次刷新，这个下载进程的时间配额也会刷新，将会继续正常运行。所以我们不会杀死这个进程，只会自动让它减慢，而用户则感受不到明显的变化，他们会发现电池消耗没有那么明显了，而程序依然在运行。

雷锋网：你如何看待无线充电技术？它会是电池问题的一个解决方案吗？

Ranveer：现在的无线充电技术主要有，感应式无线充电以及磁共振无线充电，前者需要在数厘米的距离之内充电，后者则可以在数英尺的距离内充电。我对磁共振无线充电的前景很乐观，这个技术已经被研究了数年了，目前还未能达到我们的预期，不过仍然会有很大的改进空间。

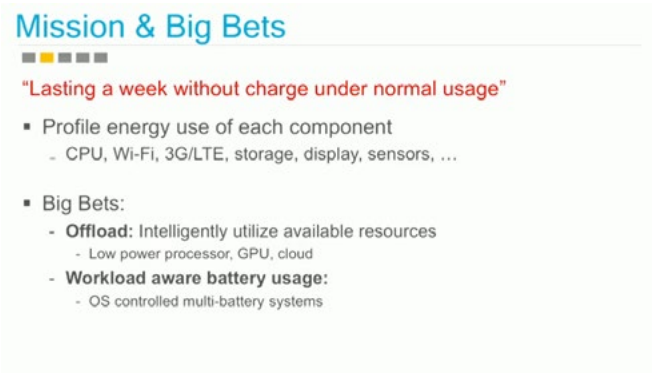
不过在目前的技术条件下，如果把充电器放在墙上，你必须离充电器在两三英尺之内；如果你离开它几英尺远，就无法进行充电了。有很多人在描绘未来的场景，即你在房间里的任何地方都能充电，现在还不行，但几年之后应该可以做到。

雷锋网：随着无线充电技术的发展，人们是否将不再对电池的电量有很高的需求？

Ranveer：我不这么认为，我想（无线充电和更好的电池）都是需要的。想象这样的场景：我在路上开车，我不知道什么时候能到达一个我能重新充电的地方……所以二者需要结合，如果你有足够多的电量，同时你又能轻易找到无线充电的地方，并且快速充电，你就不用再考虑设备电量的问题了。

雷锋网：如果电池问题解决了，未来人们的生活会变成怎样？

Ranveer：显然生活会好得多，你不用再关心充电的问题了；你的设备将总是有电的，至少电池不会再成为限制你去做你想做的事情的因素。



雷锋网：那么电池技术什么时候会迎来革命性突破呢？

Ranveer：我也想知道（笑）！一种方式是去尝试预测并等待巨大变革的发生。但今天我们可以用整合不同电池的多电池技术（multi-battery technology）来解决很多问题。即便你有了更好的电池，它也不可能满足未来所有的需求；而多电池技术则给我们提供了一种更好地使用电池的方式，你可以把它放在设备里，并立即开始使用它。

我想说的是，有很多创业公司在开发新的、很酷的电池，而多电池技术则让我们可以马上开始使用这些电池，而不是去等待更多的创新，或者更大容量、更高电流的电池。

@微软亚洲研究院 官方微博精彩选摘

#微软研究院最新技术早知道#

【手机如何变身医疗急救设备？】出行时，同行人突然休克或受伤血流不止怎么办？这时，在救护车到达前，最好能实施正确的紧急救助。微软研究院研究员Nina Mishra通过搜索引擎研究使手机变身为医疗急救设备。

手机会检测用户所执行的搜索是否为紧急搜索，如果是，手机会先提醒用户拨打急救电话，随后通过搜索引擎给出一系列的急救指示。

当紧急搜索开启时，手机会先提醒用户拨打急救电话，随后给出急救指示。首先，手机中的定位传感器可以帮助找到附近的自动体外除颤器，用户可让人帮忙去取。随后，用户将手机放在病人的胸前并开启胸外按压器，此时机内加速器会提示正确的按压值及按压速度。这些操作完成时，救护车也差不多到了。

【Jeannette:我们正置身人工智能的春天】语音助手Cortana算得上微软近年向机器学习和人工智能领域推进的重要成果。在微软全球资深副总裁Jeannette看来，与Cortana互动更像对话交流，而不是机器与人的机械问答。她说: "曾经人工智能处于寒冬时节，而现在我们将迎来它的春天。"

【对话Leslie Lamport】听新晋图灵奖获得者Leslie Lamport分享他关于分布式系统、状态机等研究的体会。他回忆了贯穿自己的整个读书研究生涯的计算机和数学，并分享了自己对研究的激情，以及对数学孜孜不倦的热爱。

【微软大数据预测世界杯】6月13日，2014巴西世界杯开幕，足球将陪伴球迷们度过这个夏天！享受看球过程，也期待最终结果——冠军终将属于谁？微软研究员David Rothschild结合基础赛事数据和预测市场数据，暂定巴西队夺冠呼声最高。但随着比赛开始及更多新数据的产生，是否会有变数？

David Rothschild在很大程度上依赖于预测市场的数据。原因很简单：纯粹基本模型的问题是，即使是最好的基本模型也缺乏数据，因为世界杯每四年才举行一次，而且没有任何常规赛。其中有很多特质是难以从历史数据集中捕捉到的。

Rothschild还解释道："与很多体育赛事不同，世界杯没有常规赛，所以每场比赛都能告诉我很多信息；而且比赛持续时间较长，意味着我可以在每场比赛之后慎重地更新，更不用说在比赛期间了。"

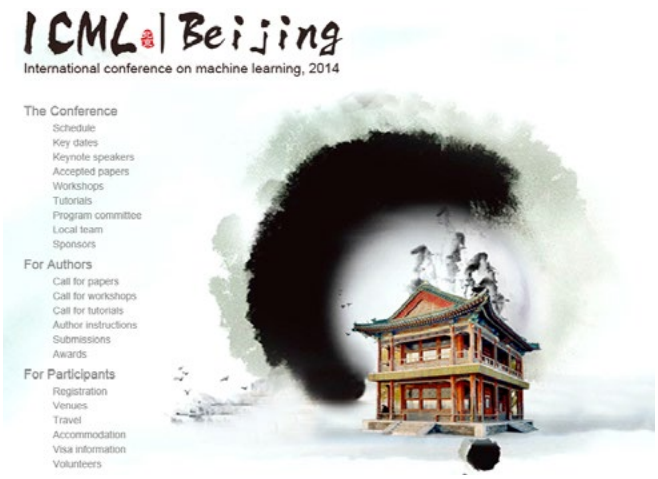
【在微软，一个社会学家能做什么？】Mary Czerwinski是微软研究院人机交互组的研究员，可她是一个社会学家，不钻研算法、也不擅长工程，能扮演什么角色呢？放弃呆在大学，而选择了企业研究院，乐趣在哪里？在她眼中，对社会学家而言，学习哪门编程语言会更得心应手一些？

Mary谈道："学习一门编程语言带来的思维改变是难以想象的。事实上，当你学会了程序员思考问题的方式时，跟他们合作时的工作效率提高了可不是一点点。"

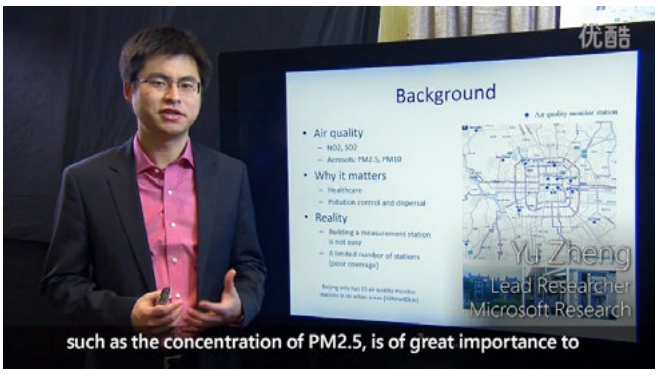
#微软亚洲研究院资讯#

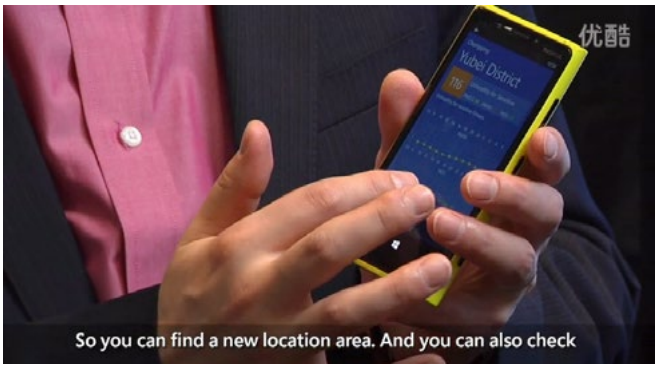
【国际机器学习盛会首次来到中国！】2014年国际机器学习大会（ICML）将于6月21-26日在国家会议中心隆重举办。本次大会由微软亚洲研究院和清华大学联手主办，是这个有着30多年历史并享誉世界的机器学习领域的盛会首次来到中国，已成功吸引海内外1200多位学者的报名参与。

往年参与国际机器学习大会的国内学者在几十人左右，而本次大会报名的国内学者超过400人。大会还邀请了微软雷德蒙研究院院长Eric Horvitz（22日），宾西法尼亚大学教授Michael Kearns（23日）以及加州大学伯克利分校教授Michael I. Jordan（24日）分别做主题演讲，值得期待！



【看大数据如何推断空气质量】PM2.5、重雾霾带来城市空气质量问题让所有人焦虑不已，然而具体地点的空气质量难以得知。微软亚洲研究院主管研究员郑宇通过构建机器学习和数据挖掘算法，实现了对具体地点空气质量的监测，并在移动设备和官方网站上提供实时、细粒度的空气质量状况。





【微软亚洲研究院新版中文网站上线】同样的域名，不同的体验!新版网站由学术合作、科技前沿、旗舰活动、新闻中心和关于我们五大板块组成，并在首页实时呈现研究院在博客、微博、视频等平台的内容，让用户更便捷地获取所需信息。欢迎登陆体验全新网站www.msra.cn，并提出宝贵意见！

【"城市计算"主题网络研讨会】微软亚洲研究院于北京时间5月28日星期三下午4：00举办"Microsoft Azure助力城市计算"网络研讨会(英文)，供世界各地同行聆听并分享观点。

这是一个不设人数限制的虚拟平台，足不出户就可参加高端技术研讨会，与城市计算领域同行、学者互动交流。该研讨会涵盖;（一）"Microsoft Azure助力城市计算"学术合作主题研究计划介绍（二）微软亚洲研究院在城市计算领域相关研究的进展（三）案例分享：利用智能手机和群智感知计算实时监测城市噪音。

Skype Translator，梦想照进现实
——微软研究院助力Skype实现跨语言的实时对话

文章译自：Enabling Cross-Lingual Conversations in Real Time

有时候，一个梦想只需要轻轻一推，就能变成现实。

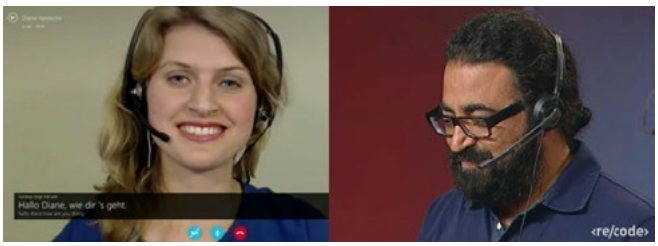
大约一年前，时任微软首席研究和战略官的Eric Rudder和微软研究院院长Peter Lee把这个理念带给了微软机器翻译团队。那时，Eric和Peter一道观摩了一个研究项目的演示，该项目旨在为使用不同语言的人提供流畅的跨语言对话。这个演示向观众展现了一个十分诱人前景——将有望突破几千年来一直阻碍人类交流的语言障碍。

看罢演示，Eric和Peter深信：实现这一夙愿的时机已经成熟。而后，在雷德蒙总部由Arul Menezes领导的机器翻译团队成为担负攻坚任务的几个团队之一，负责将这一令人振奋的研究成果转化进入微软的产品中。

推进语音翻译产品化成果初现

5月27日，微软CEO Satya Nadella在美国举办的Code大会上发言时，特别谈到了这个团队迄今为止所取得的成果，并邀请Skype副总裁Gurdeep Pall上台首次公开演示了Skype Translator应用。在演示中，说英语的Gurdeep与说德语的微软员工Diana Heinrichs在使用该应用后实现了流畅的实时对话。此次演示的成功展现了微软研究院机器翻译团队在稳步推进产品转化的过程中所取得的阶段性成果。

“我们认为，语音翻译是我们一直所从事的文字翻译的一个很自然的演进，”机器翻译团队项目经理Chris Wendt说。他一直为推动这一方面的产品孵化努力着。“这是一个激动人心的项目，”他说道：“而且，我们也清楚地认识到，在Skype中加入这项功能，使人们得以在语音翻译的帮助下进行自由交谈，将成为这一技术的杀手级应用情境，非常有助于把这项技术带到用户面前。”



过去的数十年，人们在语音翻译领域已经有过诸多尝试，其中一些就发生在微软研究院内部。但是这些尝试要同时满足能够应用在现实生活中所使用的工具中，并适应许多不同用户的嗓音和说话习惯以及不同语言之间的细微差别，还要服务于像Skype这样规模庞大的用户群，则往往都被视为是不可能完成的任务。

今年晚些时候，Skype Translator将作为限量测试版（limited beta）首次出现在Windows 8中。这就需要微软在翻译、语音识别和语言处理等方面实现一系列重大的研究进展。当然，这离不开来自全球各地的微软工程师和研究团队的共同努力。

微软研究院机器翻译团队是一支有着辉煌战绩的队伍。它曾负责研发前沿的翻译技术和云服务技术，后被广泛应用于微软一系列的产品

中。这次，它践行了 “一个微软” 的策略来应对这一挑战——合微软语音服务领域的研究人员和工程师之力。此外，为了开发一种全新的自然用户体验，该团队还与Skype的设计师和工程师建立了深入的合作伙伴关系，尤其是以Jonas Lindblom为首的原型设计团队。

Menezes说：“这个项目的有趣之处在于我们首次把这两个相当复杂的技术组合在一起，提供这种端到端的用户体验。”

训练之上，数据制胜

微软研究院专注于机器翻译课题已经长达10年之久。最初的结果来自于由微软产品所支持的知识库的翻译。后来该技术作为 “必应翻译” 的底层引擎，开放给社会公众使用，并由此带动了这项工具的发展和普及。如今，必应翻译的用户已经数以千万计，并获得了世界各地许多组织的青睐。一路上，Menezes和微软同仁们克服了系统和用户界面设计方面的许多严峻挑战，其中包括降低时延和开发视觉反馈，这使得翻译系统能够利用用户反馈来不断自我完善。

“在这一领域，数据有多好，技术就有多好，” Menezes说。“我们工作的一大重点在于扩充对这些系统进行机器学习训练所用的数据总量和种类。”

微软亚洲研究院就曾敏锐地感知到这些数据的重要性。语音组的首席研究员Frank Seide一直在使用逐步增加会话数据的时间长度的方法，并用它们来对基于模型的训练方法进行微调。开始时，只有一组总计24个小时的数据，而现在其数量已经显著增加。由Frank研发的先进的语音对话模型，在这次Code大会上在所演示的无缝交流中发挥了至关重要的作用。对话训练的数据很难获得，团队不得不开发新的技术来收集对话语音数据。Frank说，其中一个例子是对社交媒体上的发帖进行分析。

“这让我们接触到另一种语言风格，” 他说：“人们在社交媒体上的写作方式不同于他们的说话方式，但其中的俚语和词句仍然会有一些重叠，可以对系统有所帮助，并让它跟得上时代步调。” 但尽管这样，还存在一个 “不流利” 的问题：也就是人们写作和说话方式之间的差异。在说话时，人们使用大量的停顿和无意义词句，用于填充思考过程中的空白。

“人们不只会重复一个字，” Menezes说。“有时候，你的一句话只有三个字，还会回头去重复一遍。在某些语言中，这方面的挑战性更强，例如在西班牙语等语言中，用词必须与语法性别保持一致。“解读这样的交谈需要进行大量的训练。确定一个句子在何处终止也是这样。句子是翻译的基本单位，但如果没有标点符号，就可能很难辨认。事实上，如果缺乏标点符号，阅读一段对话的转录文字也是几乎不可能的事情。翻译器必须学会对语音输入进行分割。

“这就是过去一年间我的团队一直在做的事情：解决人们说话方式和写作方式之间的不匹配，” Menezes说：“如果你的翻译系统只侧重于书面文字，那么它在口语上的表现一定非常差。”

深度神经网络为语音翻译注入新生命

除了机器翻译外，语音翻译工作还要依靠其他几个方面多年来的研究和工程开发的进展，而其中最值得注意的则是语音识别。在这一领域，微软亚洲研究院和雷德蒙研究院的语音研究人员都发挥了关键作用。通向Skype Translator的道路在2010年秋天开启了高速模式。Frank和同事Kit Thambiratnam开发了一个名为 “The Translating! Telephone” 的系统，用于电话交谈中的语音到文本及语音到语音翻译。

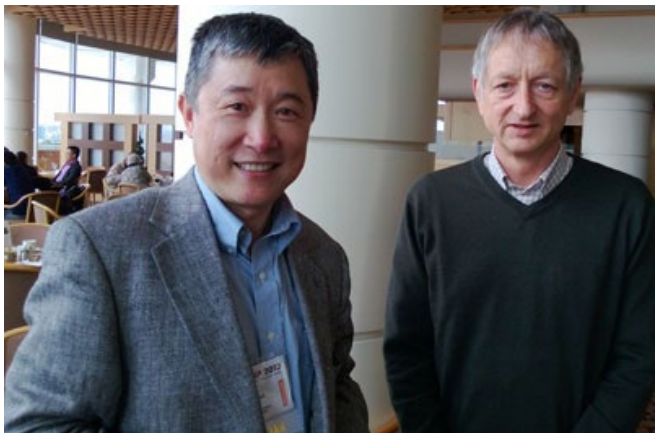
另一个重要突破来自俞栋和邓力与硕士实习生George Dahl的合作，他们使用句音和发音小片段作为神经网络的训练标的。此举让系统性能大大提升，语音识别的准确性提高了16%。“我感到非常兴奋，” 雷德蒙研究院的研究员俞栋微笑着说：“主要是因为我专注于这个领域已经有很长一段时间了，而且我知道在我的职业生涯中从来没有出现过如此大幅度的提升。” 新的性能在Frank脑海中留下了深刻的印象，并被它深深吸引，于是乎一段密切的合作由此开始。经过一番时日的努力，准确度提高了24%。进一步的调整后，最终实现了42%的改善幅度。可以说，大器已成。

这时候，Frank不得不重新设计自己的基础架构，以应对更大规模的任务。他与微软亚洲研究院的李钢合作，致力于核心语音识别技术的开发，特别是构建新的模型。每个模型都可能需要数周时间。要实现Skype Translator，这里也不得不提到微软雷德蒙研究院的研究员邓力。2009年，他邀请多伦多大学教授Geoff Hinton访问雷德蒙，共同研究新的神经网络学习方法，而研究的基础，则是Hinton和他的合作者在2006年时发布的几篇有深远意义的论文，为神经网络在信号建模和分

类的应用注入了新的生命。

究竟什么发生了改变？

原来一方面，大数据的扩散给神经网络注入了新的生命。某些蛰伏多年的机器学习方法也得到了新鲜血液。另一方面，邓力表示，“深层”也发挥了作用。在上世纪90年代，人们认为神经网络方法用于语音识别的关键，是因为它拥有巨大的隐层参数，它可以用于拟合分类结果，并改善语音识别的性能。



“早期人们并没有意识到使用多个隐层层是非常重要的，”邓力说：“但其实大脑中许多部位的结构就是这个样子的。”

2012年1月，微软的研究人员共同撰写了一篇论文，题为《Context-Dependent Pre-Trained Deep Neural Networks for Large-Vocabulary Speech Recognition》，发表在IEEE音频、语音和语言处理（IEEE Transactions on Audio, Speech, and Language Processing）期刊上。这是“深层神经网络”术语首次应用于此类工作。

如今，“深层神经网络（DNN）”这个新生术语已经在语音识别界普遍使用。四年前，IEEE声学、语音和信号处理（IEEE International Conference on Acoustics, Speech, and Signal Processing）国际会议仅有三篇关于深层神经网络的论文发表，而今年的会议发表了近百篇相关论文。那篇由俞栋、邓力等人合作的论文被引用得如此频繁，以至于最近荣获了2013 IEEE信号处理学会（IEEE Signal Processing Society）最佳论文奖，这项经得起时间检验的奖项用于表彰过去五年间发表的顶级论文。使用音符组合（senone）作为语音解码主要单位的方法还带来了其他的好处。包括微软在内的大企业都为语音处理能力的建设投入了巨资，这些架构都是在深层神经网络出现之前建成的。以前的模型都是以高斯混合模型为基础，而音符组合则是其输出形式。

DNN方法也可以将音符组合作为输出，从而实现了对现有计算架构的有效再利用。这些突破已经带来了回报——微软的必应平台产品组正在利用它们为Cortana等语音驱动的产品提供业界一流的语音识别体验；Skype、微软研究院和必应平台组之间的这种紧密伙伴关系已经让识别的精度接近人类的能力，并且成为向全球范围内的Skype用户提供翻译服务的关键所在。

梦想照进现实的注脚

对于在“一个微软”理念下开发的Skype Translator而言，一次重要的试验发生在2012年10月25日，地点在中国的天津。当时微软亚洲研究院正在举办年度学术研讨会——“21世纪的计算大会”。在时任微软研究院全球负责人的Rick Rashid发表的主题演讲中，首次公开演示了微软实时语音翻译技术。

“我要确保每个句子后面都有停顿，”Rashid回忆说：“好让观众有时间清楚地聆听中文普通话版的我说了些什么。这也意味着观众有足够的时间做出反应。我记得听到来自前排观众的大声赞叹，紧接着便是全场观众的热烈掌声和好评，这场面相当令人感动。”

对语音识别和机器翻译的发展而言，这都是一个关键的时刻。

“我认为，这次演示很明确地预示了深层神经网络语音识别和实时语音翻译的到来，”Rashid说。“它为未来的努力设置了标杆，并让我们窥见到未来：语言未必是一种障碍。我从小就是科幻小说迷。看到一些曾经纯属虚构的技术变成现实，着实令人振奋！”

现在，Wendit和他的同事们正将这一技术推广到大众中去。他率先将“一个微软”的理念运用到Skype Translator的开发中，其中包括与Skype的同事Lindblom, Daniel Nitsche, and Fredrik Furesjö的密切合作。在Skype团队的Steve Pearce以及雷德蒙研究员Shamsi Iqbal的帮助下，开发团队得以对用于用户体验的交互模型进行了微调。

“团队克服了设计、研究和工程方面的很多挑战，才走到了今天这一步。最令我感兴趣的是，引入了更自然的沟通方法之后，我们与技术之间的动态关系发生了怎样的改变，”Vikram Dendi说道。他在成为Peter Lee的技术和战略顾问之前，曾是机器翻译团队的一员。“作为一名用户，在使用传统的文本翻译时，如果没得到正确的结果，我很容易将责任归咎于技术，然后干脆放弃。”

“令我着迷的是，双方本着如此强烈的愿望尝试相互沟通，并借助技术，让彼此增进对语音翻译的理解。即使是在测试阶段，这项技术也已经变得既有用又好用了。”

IE11阅读视图：带给你静心饕餮阅读大餐的片刻

编者按：又到读书日，今天你挤出时间读书了吗？如今，越来越多人在习惯电子阅读，然而总难逃眼花缭乱的干扰信息。Internet Explorer 11新增阅读视图功能，一键开启，给你带给你静心饕餮阅读大餐的片刻！那该功能如何从技术上实现、有着怎样的研发故事？本文为你讲述这一功能从研究到产品的“前世今生”。

4月23日，我们迎来又一个“世界读书日”。

2014年的读书日里，数字化阅读和纸质阅读逐渐“平分秋色”，为不同的人群所喜爱。在这一天，请捧起你心爱的书或打开你的电脑，在温馨安静的环境里，享受阅读的快乐。

然而，困扰不少。

在用平板电脑或手机来阅读小说或跟进时事信息时，眼花缭乱的信息不断闪现，渐欲迷人眼；想专注阅读心中所爱，却难逃扑面而来的广告干扰，难以积聚心神；偷得浮生半日闲，难得有时间静心阅读，却发现文字太小且排版太密，白底背景色又过亮，搅得人瞬间兴致全无。回想起这些数字化阅读的恼人“事迹”，恐怕好多人又会不由自主地开始怀念起过去那手不释卷的日子了。

这一切的期待，都将在Internet Explorer 11（IE11）浏览器中得以实现。它新增的阅读视图（Reading View），将为你开启无干扰的阅读时刻，清除烦恼，让你享受片刻舒适沉浸式阅读的美好时光。

目前，用户可以在Windows 8.1的Modern UI界面或Windows Phone 8.1中找到IE11阅读视图功能。它能够精准地抽取众多网页的主体阅读内容，并在去除干扰信息后对其重新进行排版。当用户打开相应网页来阅读小说或关注时事时，IE11能够对该网页进行识别判断，随后生成一个阅读视图的按钮。读者点击按钮后，随即进入阅读视图。阅读视图将意味着什么呢？

- 可以暂时隐藏恼人的干扰信息、保留正文文字和图片，带来干净整洁的视图；

- 可以对网页重新排版、自适应屏幕大小（屏幕比较宽的时候，会选择横向呈现阅读内容，而屏幕比较窄的时候，会自动选择纵向呈现阅读内容）；

- 可以让读者自行调节字体、行距和背景颜色，让读者能舒适愉悦地读书。

一切的一切，就是为了让爱阅读的你，专注而舒适饕餮阅读大餐！这一新功能从技术上是如何实现的呢？它的背后又有着怎样的研发故事？

基于Windows 8.1 的Modern UI 界面

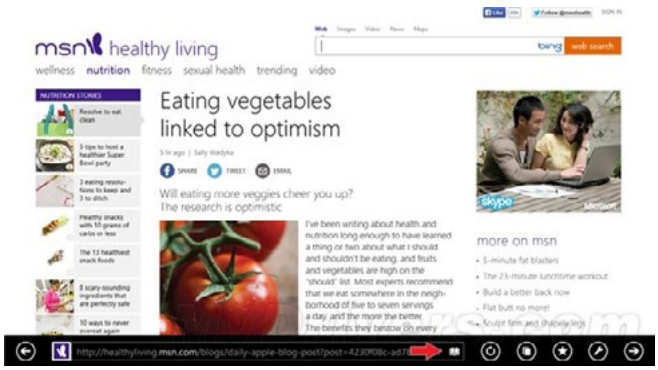


图1 初始网页

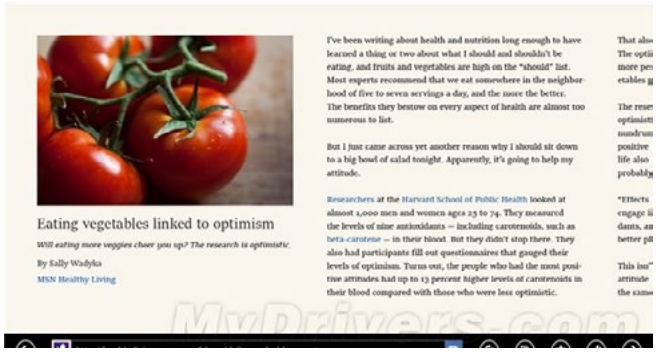


图2 使用阅读视图后的效果图

基于Windows Phone 8.1

图3 初始网页

图4 使用阅读视图后的效果图

多年的技术积淀：“网页分区”和“学习网页模版”

“这项技术，我们在11年前就开始研究了。”微软亚洲研究院主管研究员宋睿华说道。

2003年，宋睿华刚加入微软，而早在那个时候，她就已经开始接触一项为实现IE11沉浸式阅读视图功能奠定夯实基础的相关技术——“网页分区”。

当时搜索引擎才刚起步，人们在使用搜索引擎进行搜索的过程中，往往会出现不少问题。宋睿华恰好观察到一个现象，她说：“当时我在学校上学，尝试用搜索引擎搜索‘桂林三宝’。但是当时最好的搜索引擎找到的结果也有很多不准确信息。‘桂林’二字有可能是网页中间正文的名字，而‘三宝’则在右边更多推荐信息侧栏中。因此，右侧是在说吉祥三宝，而中间是关于桂林的新闻，完全不是我要搜索的内容。”

究其原因，睿华有了这样的思考，“当机器在进行搜索时，把网页的整个页面当成一个整体，没有对网页进行分区，只关注关键词是否匹配，但这明显是不准确的。当时大多数搜索引擎都是这样，它们对网站的了解和定义比较浅显，搜索时会出现不少问题。”

“人们浏览网站，会很一目了然——这个是正文、那个是横幅广告、这个是指向其他网页的链接、那个是页脚——人是很容易完成这个辨别过程的。然而对机器来说并不容易，很多时候人工智能、机器学习的技术，就是让机器逐渐具有人的能力，这就是我们在做的事情。”睿华补充道。

那如何让机器迅速地把网站分成一些区域，并知晓这些区域的功能，比如能区分正文和广告区域？

不久后，睿华离开学校并进入微软，与同事和实习生进行合作发表了一篇文章Learning Block Importance Models for Web Pages，研究对网页进行分区的重要性，以及如何让机器把网页划分成不同分区并判定每个分区的重要程度，获得了国际万维网大会（WWW）的最佳论文提名奖。后来，这项技术在必应搜索团队中初步实现了技术到产品的转化。

随后，基于睿华的工作，微软亚洲研究院发表了一系列关于网页分区的相关论文，将这一成果应用在提高搜索引擎的相关性和重要性计算中，以及如何在移动设备上更有效的浏览网页。今天看来，这些工作都是很有前瞻性的。

第二个相关技术是让机器学习网站模板。“那时候是07年了，前期研究工作是比较普遍性的，可以适用于多数网站。但如果希望做到产品转化标准，还需要进一步改进学习模板。当用户使用这项技术来抽取网页里的文字时，时常会出现很多问题——普遍性模板的抽取成功率只有85%，那剩下的15%将会给用户带来困扰。那我们如何从技术层面来补救呢？最佳方案是普遍性加上针对特定网站进行网页模板学习。”睿华介绍道。

许多网站的网页形式都非常相似，模板差不多，看上去都很一致。比如你去看央视的网站，在看了网页html来源之后，你会找到

网页布局分区的一些特点。

而睿华和她的团队做的事情与网页设计相反，是一个逆向的过程。“考虑到同一网站的不同页面具有相似性，我们要如何从不同网页中学习出其中相似的东西，这有点像归纳，归纳出一个相同的模板。而归纳模板的目标是，当机器学习了这个网站的10个页面，接下来可以快速准确地抽取来自这个网站的其他几千甚至几万页。模板技术是为了提高一些大网站的抽取精准度，我们在这其中做了很多研究。”

从技术到产品：柳暗花明又一村

文本抽取相当于前两个技术的实践应用，睿华在07年写论文时就做过文本抽取，但当时没有了解到技术转化成产品的需求，就沉寂了一阵子。直到12年12月，微软产品团队找到微软亚洲研究院，寻求合作。然而，初期的技术转化并不那么得心应手。

“07年我们发表论文时，已经给前两项技术开发了一个文本抽取的应用，当时我们认为已经做得很好了，论文被人工智能（AI）最好的会议接收了。但是直到12年微软产品团队找我们合作，我们尝试了原来的抽取模型，但发现效果不够理想，甚至离当时市面上的其他应用还差很远。”睿华回忆道。

有一段时间，睿华和她的团队成员窦志成相当疑惑，到底是出了什么问题，让一个以前评价也不错的应用变得不太好用了？

不久后，他们找到了两大原因。

一方面，与07年相比，12年的互联网网页设计改变了很多。“07年，网站设计比较单一，多利用表格来安排网页的布局。”睿华继续解释，“但2012年时，网页的结构和书写习惯跟以前已经大不同。现在内容技术先进，采取了更灵活、更复杂的技术，使网页布局看起来跟以前没有什么区别，但你打开html来源，你会发现变化很大。所以，原先老的算法就不是完全适用。”

另一方面，做研究和做产品的目标存在一定区别，在寻找问题解决方案时会出现不同偏向。睿华对此有深刻体会，“做研究，研究者希望提出一个相对确定的问题，然后寻找解决方案，这个解决方案相对比较新颖、是以前没有人提到的。但是这个东西更像一个玩具模型，你不能靠玩具积木去做一个楼，当你真的在盖一幢大厦时，遇见的现实问题就多了很多，而且解决问题的思路也不大一样。在做产品的过程中，目标只有一个——如何达到产品标准以满足用户的需求。”

在与微软产品团队合作的过程中，睿华几乎把原先论文的方案都改了，重新写了一个算法，充分考虑了产品的需求。“研究只要求我找到一个问题并解决就好了，不过做产品的话，还要考虑以后是不是便于去调整修改。从做产品的角度考虑，我选择放弃了原来的算法，重新寻找新的解决方法。”她继续说道，“其实，有时候在产品中采用的不一定是‘the state of art’的技术，我不是技术迷，我不追求一定要用最华丽的最难的技术来解问题，但我希望它是一个长期有效的解决方案。”

自此，算法上的问题基本解决。

随后，睿华的团队开始与微软亚洲研究院创新工程组的高光平、刘欣、李俊杰进行技术产品转化，合作开发阅读视图功能。经过多番讨论和落实，最终实现了在IE11中的部署——IE11的阅读视图功能对新闻或小说进行文字抽取的准确率很高，同时可以完整抽取图片，让用户根据自身爱好自行先选择文字、行距和背景颜色，不打断用户阅读的好心情！

群策群力：舒心的合作

其实最初，微软产品团队并不知晓睿华及其团队的研究，他们打算通过第三方公司的专利来实现IE11中的阅读视图功能，而微软内部的良好合作机制让微软研究院得以把睿华团队的技术分享给微软产品团队。

“正如洪院长曾讲过的，当微软产品团队有需求时，他们可以直接使用储存在微软内部的技术，这就是研究院存在的一个重要价值。在这个技术上，我们研究了11年之久，就是期待某一天，当微软需要以用户的需求为导向来优化产品功能时，我们有储备好的、现成的技术可以来直接用；相反，如果在发现产品团队对这个技术有需求时才去研究技术的话，就有些晚了，周期会非常长。”回想起这次合作，睿华这样谈道。

睿华对创新工程组的工程师们评价很高：“跟他们合作非常舒服，我主要专注在算法上，光平则会专注解决技术转化的工作，帮助在IE11的编码上实现我的算法、转化成产品功能。光平有跟Windows产品团队合作的丰富经验，效率很高。整个过程，我都很愉快。”

对于这次和睿华开展的合作，微软产品团队的首席项目经理Jane Liles评价道：“微软研究院的研究工作使我们在很短的时间内成功地在IE中实现了一项全新的功能，而不是借助第三方公司的帮助。研究院在这个合作项目上关注了大大小小、几乎所有的产品细节，极大的帮助扫清了功能上线时所必须要保证的精度问题。这是一次非常愉快并且成功的合作。”

最后

合作远没有结束，产品功能的进一步改进仍在继续，仍需要研究员和工程师的通力合作。一款优秀的产品就像一只能在蓝天自由飞翔的鸟。研究员和工程师就像鸟的两个翅膀，缺一个，鸟都很难飞起来。

你说英语够地道吗？
——必应词典“我爱说英语”来给你准确打个分

你的口语够地道吗？和美国人还有哪些距离？提升的空间在何处？必应词典“我爱说英语”将为你一一指出。

由微软亚洲研究院和微软（亚洲）互联网工程院（以下称为必应产品团队）联手打造的“我爱说英语”是必应词典的全新模块，在听说以及跟读的过程中，帮助使用者评估并不断提升自己的口语水平，最终说一口流利标准的英语。该模块现已实现了PC，Windows Phone、Android和iOS平台的覆盖。在各个平台的手机商城下载最新版“必应词典”应用，点开“我爱说英语”，一天一天练习，一口流利的英语指日可待！

“我爱说英语”如何帮助英语学习者学习英语口语？其工作原理是什么？

用户依据自己的英文水平，选择不同等级的语境。而后“我爱说英语”将会根据用户的选择给出相应的英文句子，并提供语音领读。用户在听完领读后跟读，随后系统会评估每个单词的发音，给出分数并提示需要改善的部分。“通过‘我爱说英语’，我们希望用户就像每天都随身带着一名能说地道英语的老师一样。”来自微软（亚洲）互联网工程院的首席研发总监邹欣介绍道。通过与微软亚洲研究院有着多年语音研究经验的资深研究员联手，“我爱说英语”最显著的优势是机器打分的准确率高，能真实反映跟读者的英语口语水平，因而用户能知道他的口语发音的水平和问题所在。那这一强大功能背后有什么样的研究和开发故事呢？什么样的算法支撑着高准确率的打分系统？什么样的系统开发让用户享受更顺畅、舒适的使用体验？

深层神经网络（DNN）加速语音算法研究

语音识别是教电脑怎么“听”，语音合成是教电脑怎么“说”——虽然“我爱说英语”是教语言学习者说上一口流利的英语，但技术原理都是一样，先得教电脑识别语音。

必应词典的这个功能起源于微软亚洲研究院语音组主管研究员钱瑶所负责的一项研究项目——“用深度学习的方法训练并评估发音”。介绍到这个项目时，钱瑶说道：“教电脑识别语言，必然需要很多的训练数据，随后根据这些数据训练出一个统计模型，计算机就可以自动实现语音识别。在这个项目里，我们也需要建立这样一个模型，其数据主要来源是两百多个美国人的语音数据，都是标准地道的发音数据集”，“最近比较火的机器学习方法是深层神经网络，我们也使用了这个方法，学习出一些模式（pattern），让计算机在听了跟读者的发音后自动做出判断，告诉你哪些音发得好、哪些音发得不好。”钱瑶继续介绍道。

那么，为什么基于此方法，模型的打分系统准确率特别高？“语音模型的学习在深层神经网络（DNN）阶段得到了很大的突破——深层神经网络（DNN）对数据建模、尤其是时变的函数建模、随机程序建模是很有效的。而语音是其中的典型例子，比如一句话就是一个时变函数，能量强度随着时间和频率而变化。”微软研究院语音组首席研究员宋謂平解释道。

“在研究中，我们需要找到这个深层神经网络中的输入和输出的关系。在输入是有很多特征函数的语音后，输出对于某一些特别的语音单位的后验概率。利用这种映射，我们可以找到他们之间的关系，这是一个随机函数的概率模型。对于信号本身来讲，它要求的假设很少。使用深层神经网络的学习方法，基本上你只要找到输入和输出的对应关系，它就能训练出一个模型。而相比之下，以前所使用的高斯混合模型（Gaussian mixture models， GMMs）准确性的确不够高。”宋謂平进一步解释道。

此外，计算机是把跟读者的发音与数据集中的两百个美国人的发音样本进行比较，考虑多种语音音色、声调的可能性；并且在计算计较过程中，按照音素（甚至分成比音素更小的次音素）来将用户的发音与标准发音“对号入座”，检查差异，提供准确的发音得分。“我们的数据集里有两百多个人，其中包括男男女女、老老少少，语音识别基本上也是一个概率统计分布的问题。经过比较后，你的声音（包括语音、语调、停顿、频率等）跟这两百个标准美语人当中的任何一个人较接近，那你的得分就会比较高。”钱瑶解释道。

从算法到产品

算法到产品之间的距离要如何填补——如何把研究仅需的桌面单机程序转化为用户能舒心使用的产品。这时微软亚洲研究院创新

工程组的同事黄贤俊和郭旭就开始扮演重要的角色。他们在研究员与必应产品组开发工程师之间架设起一座桥梁，扫除各类障碍，最终实现功能的上线。

深层神经网络（DNN）学习方法带来极高的准确率，却也带来了一些问题。其中最主要的问题是该模型所占用的空间很大，运行速度也慢得多。对于产品而言，这是大忌，既消耗和占用许多有效资源，也不能给用户带来及时反应便捷、舒适的体验。必应词典中的“我爱说英语”模块的语音评估服务基于微软公有云Microsoft Azure平台运行，如果不能有效解决这个问题，产品化将难以继续进行。

为实现最终的产品化，微软亚洲研究院的研究员和工程师做了不少工作，包括缩小模型、减少其占用的空间，对其进行提速。“我们使用了奇异值分解（SVD, singular value decomposition），把无效没用的东西进行压缩，在不损失效用的前提下压缩模型的数据。这样，模型数据的正确性几乎没有损失，其大小却适合上线以提供用户服务。”宋謌平说道。“从算法到系统的开发过程是重复迭代的，不断发酵、优化，是螺旋式反复上升的过程。”微软亚洲研究院创新工程组的工程师黄贤俊谈道。在后期的测试中，必应团队邀请了10位参加托福考试的同学参加产品原型的测试，与他们的托福口语成绩进行比对，检测打分准确率的高低。

而最后，他们的托福口语实际成绩与“我爱说英语”打分结果的相关性超过80%。其实早在2012年，在微软雷德蒙总部举办的微软技术节上，宋謌平和创新工程组同事郭旭合作演示了这项技术成果。在展示中，他们发现——与多数土生土长的中国人相比，外国人或有海外经历的华人的口语得分明显高出一大截，二者有质的区别。这也侧面表明“我爱说英语”打分系统的准确率是很高的。

更多的期待

谈道这次合作，宋謌平笑着说道：“这是一次很愉快的合作。其实，这不是微软亚洲研究院和必应产品团队的第一次合作了，频繁的沟通让产品转化的想法得以发酵，而多次的合作又让转化的过程更为流畅，已经是轻车熟路，顺水推舟了。”

微软（亚洲）互联网工程院的首席研发总监邹欣对此也深有感触。他曾经在研究院带队七年，负责产品转化和创新。他说：“许多创新的想法都需要快速在实践中收集反馈，产品和研究部门需要紧密结合，不断琢磨如何把技术的新突破通过合适的软件工程的手段和设计，给用户带来实际的价值. 我们都期待把更多前沿研究成果带到用户面前。”

必应产品团队的工程师李春旭也谈了谈自己的感受：“虽然在开发初期，供用户学习发音的素材还不够丰富，但打分的准确性却非常高，让我们对开发充满信心。此外，随着访问量的上升，产品代码也相对稳定可靠。如果没有微软亚洲研究院的研究员和工程师的热心和持续合作，我们很难做出这么棒的产品！”

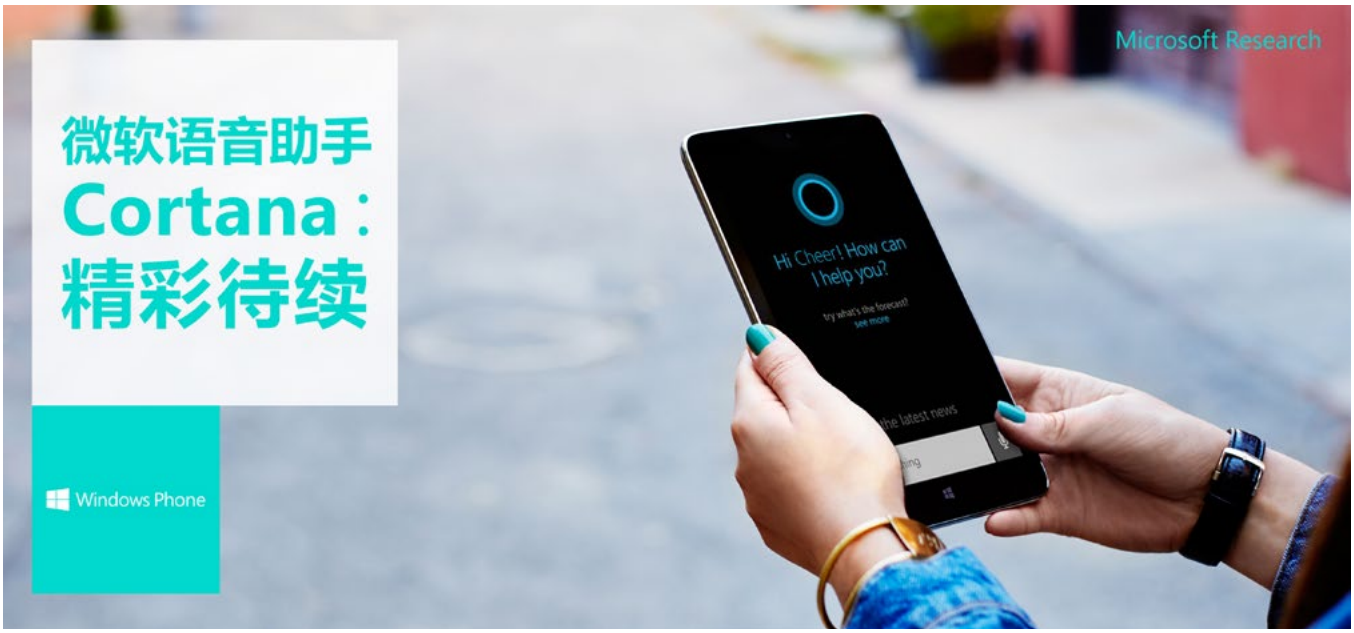
“必应团队大约花了2周左右的时间就开发出了第一个版本，还是比较快的。并且，在开发过程中，加入了互动、社交的元素，比如可以分享测试结果等，让产品变得更有意思、更有乐趣。” 微软亚洲研究院创新工程组的黄贤俊谈道。贤俊对教育类应用有着不少期待，他说道：“去年，我开始接触教育类应用，一直想做所有人都能使用的好应用，真希望‘我爱说英语’能给更多英语学习者带来好处。”

未来，“我爱说英语”的功能将不断完善，带给用户更多的惊喜。期待科技与教育的结合，将惠及更多人！

Cortana：精彩待续

文章译自：Anticipating More from Cortana

编者按：借助必应搜索引擎赋予的力量，微软语音助手Cortana能够模拟人的语气及思考方式与用户互动，提供更多帮助。Cortana包含微软研究院在多个研究领域的成果，博客中微软杰出工程师Larry Heck分享了Cortana从基础技术到产品体验的过程，以及更长远的目标。



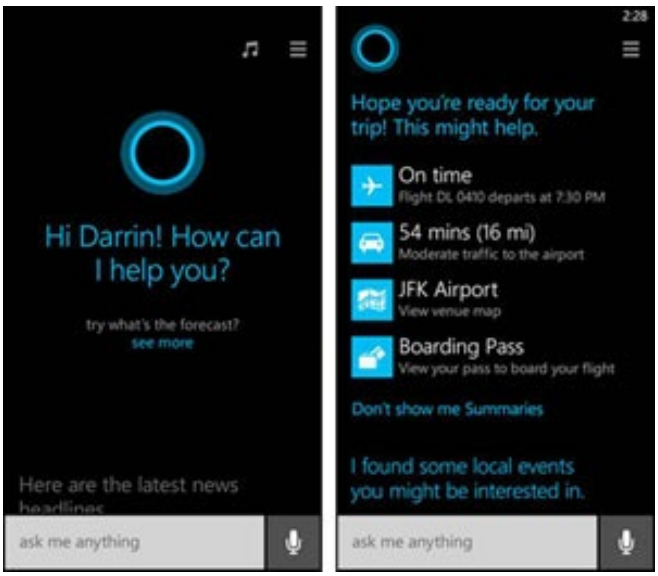
我们大多数人都梦想着拥有完美的个人助手——需要的时候它总会在那里，能够时刻读懂我们的心，知晓我们的每一个需求，并悄悄地帮我们安排生活琐事。伴随着Windows Phone 8.1的推出，全新数字个人助理Cortana终于揭开了神秘的面纱。它被必应搜索赋予了强大的力量，让用户离梦想又近了一步。

对于微软研究院杰出工程师Larry Heck而言，Cortana首个版本的发布让他尝到了心中所想变为现实的甜头。展望未来，Heck期待Cortana能够越来越善于推测用户的需求，并且以更加自然的方式进行互动。事实上，Cortana已经在某种程度上实现了这一点。它不仅能简单地执行声控命令，还会不断了解用户并且逐渐实现个性化，这样一来Cortana就能主动地在对的时间执行相应的任务。如果用户每天下午离开办公室前都会询问室外温度，Cortana就能学会无需询问便自主提供相关信息。

此外，如果Cortana有权限访问手机数据，就可以读取日历、联系人和电子邮件信息，更好地理解用户所处情境和人脉关系。Heck在当地乐团演奏古典小号，他可能会收到有关排练时间变动的日历更新。Cortana会告知他时间有变，而如果新时间与另一项约会冲突，它也会发出提醒。

优势来源于研究深度与广度

虽然很多人会将此类逻辑关联和类似人类的行为笼统地归类为“人工智能”（AI），但Heck指出，研究领域的多样性对Cortana的底层技术做出了贡献。他认为，Cortana是微软研究院在与个人助理技术相关的多个领域工作的具体体现。



“虚拟个人助理所需的基础技术包括语音识别、语义/自然语言处理、人与机器对话建模以及口语生成，”他说：“每个领域中都有一些微软研究院多年来一直试图解决的研究课题。事实上，在其中的每一个领域，我们都是先行者。”因此，Cortana的设计理念是植根于最先进的机器学习和数据挖掘算法。不仅如此，开发人员和研究人员都能够使用微软横亘商业和企业产品的广泛资产，包括它们与必应搜索和微软语音算法及数据之间的牢固关系。

如果说Heck为Cortana的未来设定了高标准、严要求，那么这正是微软研究院内部深厚且多样的专业知识积累的结果。“微软研究院在人工智能领域有着悠久而广泛的经验，”他说：“这里有人工智能领域领先的科学家和先驱者。这项工作的愿景和努力方向源自Eric Horvitz对会话互动的研究和理解，这可以追溯到90年代初。语音和自然语言处理是长期存在的研究领域，机器学习也是如此。此外，微软研究院还是深度学习和深度神经网络研究的领导者。”

从基础技术到整体体验

2009年，Heck在微软开始了当时被称作“对话理解”（CU）个人助理的研究工作。“当时我在Satya Nadella领导下的必应研发团队工作，” Heck说：“这项工作的技术愿景是开发虚拟个人助理。不久后Steve Ballmer任命Zig Serafin来统一领导微软公司上下的各种语音研究工作，Zig邀请我加入团队并担任首席科学家。在这个岗位上，我与Zig合作，制定了详细的计划，并开发出现在人们所知的Cortana。”

曾共同开发Cortana产品的研究人员：（左起）上排-Malcolm Slaney、Lisa Stifelman和Larry Heck；下排-Gokhan Tur、Dilek Hakkani-Tür和Andreas Stolcke。

Heck和Serafin在将对话扩展到网络广度的基础上，为微软的数字个人助理技术制定了愿景、使命和长远规划；他们还建立了一支具有专业知识的团队，来打造Cortana的初始原型。作为项目启动的第一步，Heck的团队为产品团队招募并培养了几名博士级工程师来从事开发工作。

“正是由于搜索和语音技能的组合是独一无二的，” Heck说：“我们需要确保微软拥有对的人才和技能来实施这个项目，于是我们聘请了最棒的人来做这件事。”团队就位之后，Heck和他的同事加入了微软研究院，继续立足长远，致力于研发下一代个人助理。参与这些早期工作的主要研究人员包括微软研究院高级研究员Dilek Hakkani-Tür和Gokhan Tur以及首席研究员 Andreas Stolcke。Heck团队其他早期成员包括首席研究软件开发员Madhu Chinthakunta和首席用户体验设计师Lisa Stifelman.。

“我们从低层次的基础技术入手，” Heck回忆道：“接近项目结束时，我们团队进行了高层次、全方位的可用性研究，为产品组提供了指引。就像爬到轮船的瞭望台上俯视整个体验。微软研究院的研究经理Geoff Zweig负责主导可用性研究。他带人过来试用了原型——让他们拿起来就用。然后，我们从中汲取教训。微软研究院是研究可用性的好地方，因为我们既了解基础技术，又懂得长期愿景和工作方法。”

更长远的眼光

自Cortana团队成立以来，Heck一直是不可或缺的组成部分，但即使在2009年加入微软之前，他就已经对CU个人助理的早期研究做出了贡献。20世纪90年代在斯坦福国际研究院（SRI）供职时，他就从事过一些深度学习和深度神经网络技术的早期研究。

Heck还曾作为SRI团队成员，为美国政府国防高级研究项目署资助的CALO人工智能项目奠定了基础。该项目旨在打造新一代的认知型助理，可以从既往经验中学习，并且在不明情况下进行智能推理。后来他在Nuance和雅虎的工作经历又让他在相关研究领域获得了更丰富的积累，而这些领域又对Cortana的强大功能做出了重要贡献。

为此他将目光投向了更远的地方。

“我认为，现在的个人助理技术好比初期的搜索技术，”他说：“它将会在有可能覆盖到的各个领域内不断拓展。历史的演进总是相似的。在90年代中期，早在搜索出现之前，曾经出现了雅虎目录。由于它能够组织信息而很流行，但随着网络的成长，目录模型变得尾大不掉。就在这个时候，搜索出现了——现在你可以在网络上搜索任何东西。”

Heck表示，个人助手将会沿着类似的轨迹前行。当前主要致力于实现最常用的功能，如提醒和日历等，但随着技术的成熟，个人助理必须扩展到其他领域，使用户可以随时随地获取任何信息，并处理任何事务。

“微软开发了Cortana，就是要扩展到所有不同领域，” Heck说：“有长远的目标意味着我们有长期的架构。我们的目标是支持所有类型的人机交互——无论是语音、文字或手势——跨越信息和功能的各个领域，并且让自然交谈变得更加容易。”

破译手势在对话中的意义

文章译自：Kinect for Windows Helps Decode The Role of Hand Gestures during Conversations

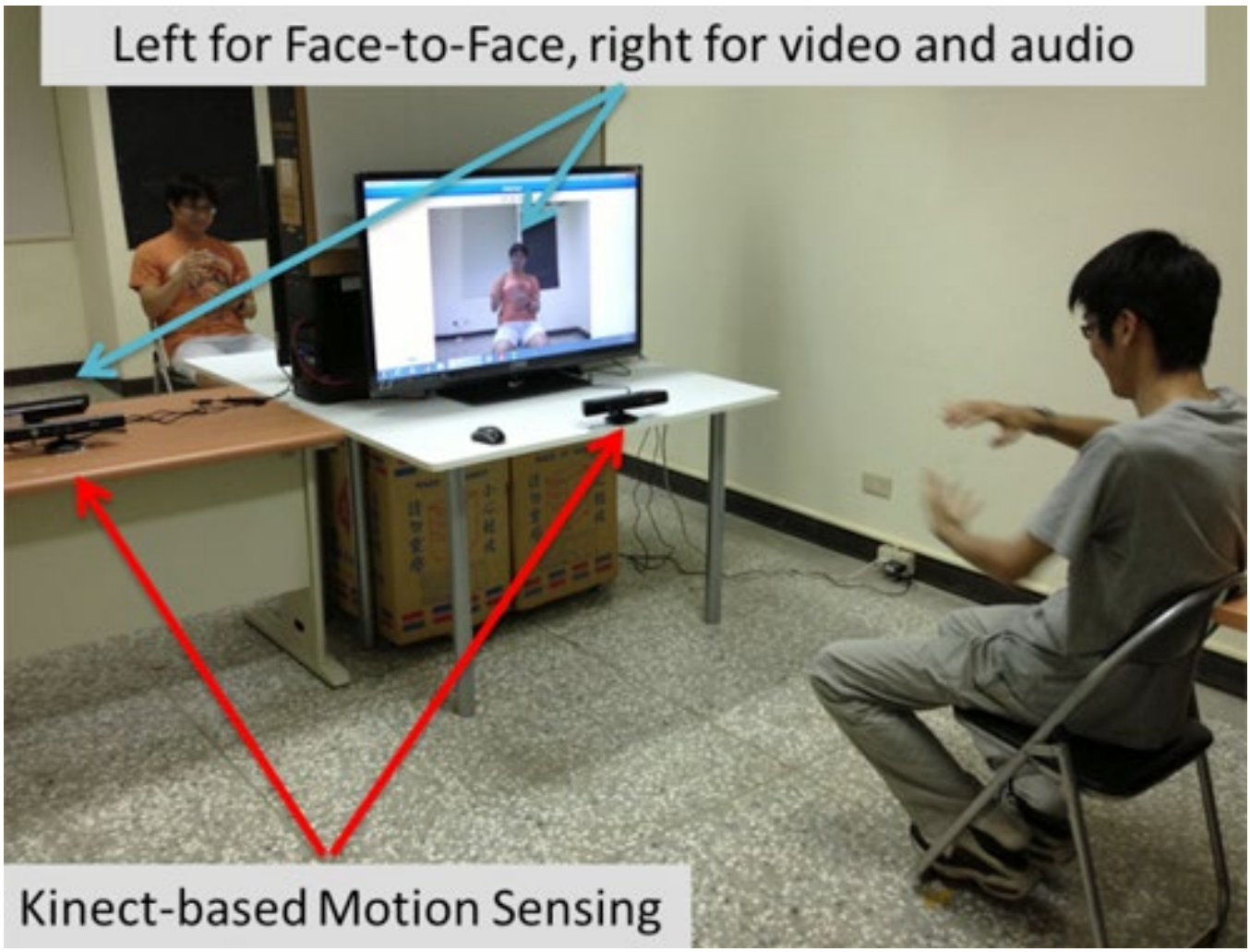


图1 实际实验场景，Kinect传感器用来支持面对面“视频通讯”的情况

我们都知道除了语言外，人类还有很多其他的交流方式——例如生气的一瞥或是勉强地点头。但除了这些明显的肢体语言，我们在交流中还频繁地使用我们的双手。对话中的手势虽然无处不在，但分析它仍然很难。我们很难得知，这些伴随着谈话的自发的手部动作是否、以及如何塑造沟通的过程和结果。行为科学家想了解这些非语言交际行为的作用。基于这个情况，技术开发者也同样渴望建立一个帮助人们更为顺畅地交流、理解信息的工具。

为了破解和他人说话时发言者双手动作的潜在含义，研究人员需要以高效且经济的方式追踪分析手部的动作路径。台湾清华大学的王浩全教授和他的团队意识到他们可以通过使用Kinect for Windows传感器去捕捉记录面对面交流过程中的手势和语音，从而解决这个问题。

“我们之所以想用Kinect，是因为它是市面上最流行和实用的运动传感器。这种普及度可以增加我们研究方法的潜在影响力。”王教授解释道。“其他研究人员能够比较容易地沿用我们的研究方法，或是重复我们的研究。Kinect的使用也让大型行为学研究成为可能，只要用户使用Kinect，我们就能远程收集他的行为数据。此外Kinect的软件开发工具包也易于使用。”

基于Kinect for Windows的诸多优势，王教授与微软亚洲研究院合作，使用Kinect for Windows在对话中捕捉观察对象的手部动作。“我知道微软研究院在应用Kinect进行一些先进的研究，因此我们和微软的研究人员进行了愉快的合作，来为学术界和社会提供有趣的研究成果。”

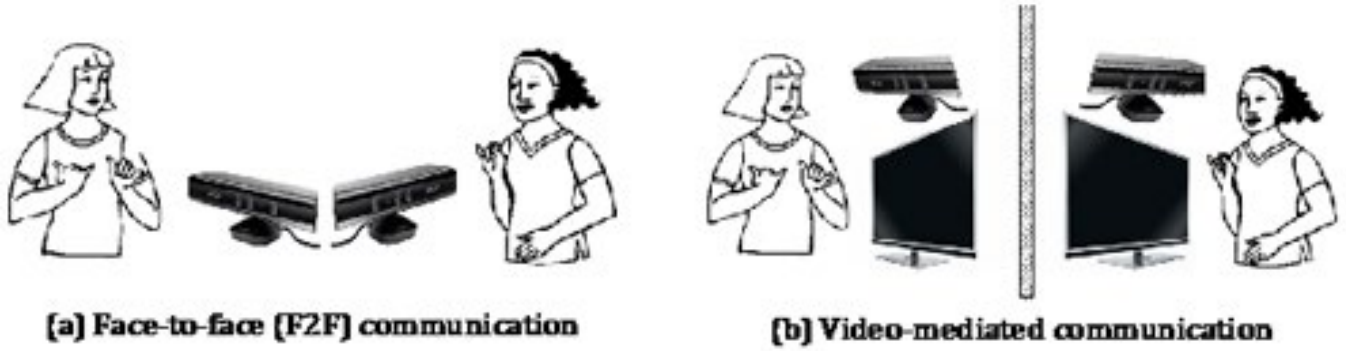


图2 实验期间Kinect Windows传感器位置示意图

在接下来的合作研究中，团队在两位对话者中间背靠背地放置了两个Kinect传感器来记录谈话过程。传感器同时捕捉到两个对话者的语言和手部活动，从而提供一份交流双方的言语信息和手部运动轨迹的带时间标记的记录。

为了展现研究方法的有效性，研究者将面对面交谈、视频聊天和音频聊天三种情况下的手部动作的数量和相似度做了对比。两个参与者在面对面交谈和视频聊天中可以看到彼此，但在音频聊天中他们看不到彼此。

研究人员由此发现了一些意想不到的结果。“第一，也许和直觉相悖，实际上人们在音频、面对面和视频聊天中的手势是一样多的。这表明人们之所以做出手势不是为了别人，而是为了自己，尽管对方看不到自己，但仍然会做出手势，”王教授说。

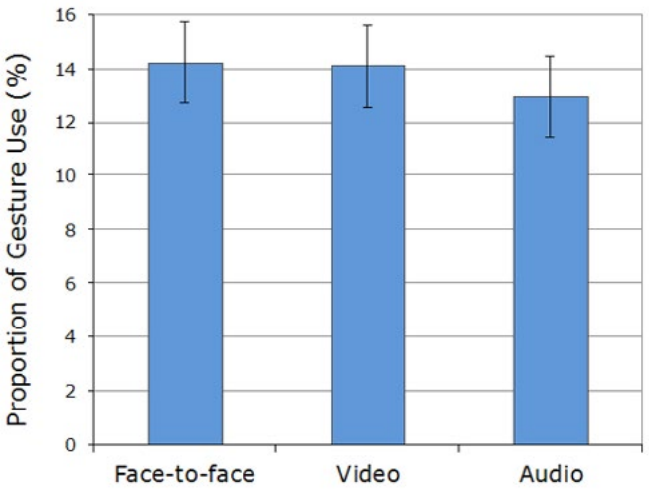


图3 图表显示，“边聊边做出手势”在只有音频的谈话与其他方式中有着相同的比例，即做出手势的人主要以自己的意志为主

研究者还考察了对话的参与者对整个交流的理解与使用手势的数量之间的关系。出人意料的是，他们发现参与者的理解水平和他的对话者使用的手势数量之间并无相关性。相反，研究者发现，参与者对对话的理解程度与其自身的手势数量相关。这从另一个角度证明，手部动作并不是为了产生和传递信息，而是反映了说话人在内容理解上的自我强化和自我认同。

无论从理论到实际，Kinect传感器都有着广泛的潜在应用（例如，当人们在交流时，我们是否可以用Skype来捕捉他们的手势动作）。此外，研究人员终于可以突破以往的速度和规模限制来设计他们的实验了。

“在进行跨学科的研究时，Kinect操作简单、有效，因而降低了研究成本——不需要把时间浪费在一些技术问题上，而能专注解决研究问题。”国立清华大学王教授谈道。

今年4月在加拿大多伦多举办了主题为“计算系统中的人为因素”的ACM SIGCHI会议（the ACM SIGCHI Conference on Human Factors in Computing Systems）。会上发表了王教授与微软亚洲研究院这一合作项目的论文全文。

“与微软亚洲研究院合作十分愉快。在感兴趣的课题上我得到了很大的支持和自由。这使协作变得真正独特并具有价值，” 王教授说，“我希望与微软的研究人员们有更为密切的合作，同时扩大当前的工作。能加强对一些非常规交流的理解，如跨文化、跨语言交流、师生授课等。因为基于语言的沟通在这些情况下经常效果不好，非语言部分也许会让交流变得更加有效。对过程的更深层的理解可以让技术设计更好地支持这些情况。”

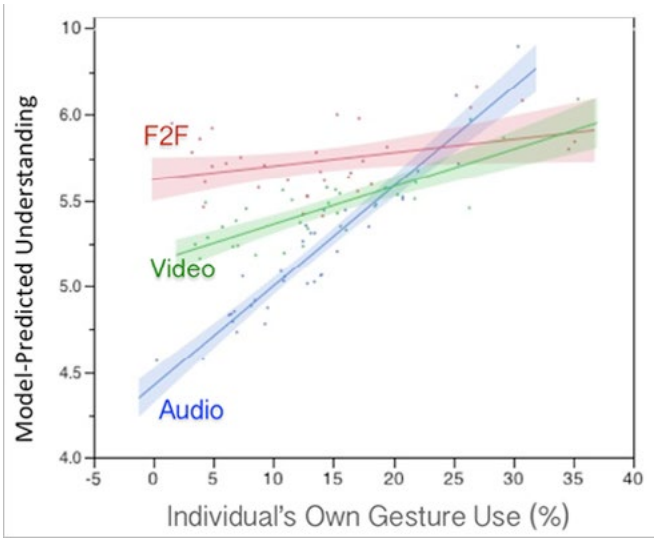


图4 自己的手势和自我强化理解程度之间的函数关系

科研不是比赛，而是一种对未知和完美的自我追求
——跟邢波（Eric Xing）面对面聊科研

编者按：6月26日，2014年国际机器学习大会（ICML）在北京国际会议中心完美落幕。作为机器学习领域两大顶尖年会之一，这是ICML大会30多年来首次来到中国和远东，在国内的机器学习界震动不小。身为本次大会主席的卡耐基梅隆大学计算机系教授邢波（Eric Xing）为此做了诸多努力。作为在美国机器学习领域前沿为数不多的华人学者，许多国内计算机专业学生渴望投身其门下。借着此次大会的举办，我们也有幸和Eric坐下来，面对面的聊一聊他对科研的态度以及对后辈们的建议。

问：从生物学博士到计算机科学博士，再到机器学习领域的佼佼者，这样的转变是如何发生的？

Eric：为什么会选择进入一个新的专业？因为兴趣。当时对手头做的研究项目兴趣不够，而同时又有其他的东西让自己很着迷，所以在完成前项工作后就选择了转变。现在回头看，这样的转变在我的生活中并不少见，比如，我最近开始在做计算机系统相关的研究。操作系统跟机器学习是很不一样的领域，这种转变的距离和我之前的转变一样远，但兴趣使然。15年前，人工智能和机器学习在美国学术圈很冷门，中国人不多、美国人也不多，当时基于兴趣选择转专业，现在也一样。当选择改变时，不要把它当个大事，要把生理和心理障碍压到最低。

问：不少人认为，做科研极为枯燥。你认为这里面有误解吗？

Eric：确实有人觉得科研枯燥、与产品的结合不够紧密，像是对未来的虚幻想象。这其中有一些误解。很多人基于片面的理解或有限的经验，简单的把理论研究和应用研究对立起来，并由此产生一种肤浅的排他思想；比如有些“基础研究”学者瞧不起“应用研究”学者，认为后者不严格正统，或者有些业界的工程师程序员也鄙视大学里的研究者，认为后者不实用。抱有这种想法的人很难成为领先的学者和工程师。其实计算机科研是最不枯燥的研究，优秀的计算机研究往往体现着理论和现实问题的紧密结合。它不是纯粹推导公式、设计模型，证明定理；也不是埋头编程，调参，试错。严谨的计算机科研需要跟数学打交道，这可能对于某些人来说，有些枯燥；但对于另一些人来说，这仍然很有趣并令人兴奋。事实上，谷歌的搜索引擎、微软的Kinect等都是由基础研究衍生出来的产品——当你把研究跟实际应用结合起来，兴趣是很容易产生的。或许研究过程中会比较理论，但我们是基于现实应用问题来提出研究问题。当你希望用一个漂亮的方法（正规、严密，普适，可独立重复的方法）解决问题时，你自然会使用数学手段；但是最后的实现、评估、证明等是从理论和实际应用两个方面来出发，这样就会很有意思。

问：有传闻说您每天只睡4-5个小时，做科研这么忙吗？

Eric：我不是为了忙而忙，而是顺其自然。就好像弹琴有的地方弹得快了后，慢了自己都不好意思，快反而是一种自然的节奏。有时候会更慢一些、有时候会更快一些，也没有刻意去保持。当然，其实从生活质量来说，也没有那么可怕。我也有进行规律的健身锻炼，业余爱好，跟家人一起玩，等等，我不提倡把自己弄得很憔悴、不顾家庭或不跟朋友交往。如何做到呢？不要浪费时间。睡几个小时是个人生理特征，但是在不睡觉的时间里，要怎么用？大部分人在大部分时间处于什么都没做的状态，他们既没有做公事，也

没有做私事，但我基本上不存在这样的时间，我要不就做一些跟工作有关的事，要不就抽空锻炼、弹琴，听音乐，看书，或跟家人一起活动。虽然没有一个明显的工作到生活的某个切换点，但找准了工作和生活的节奏，还是很自然。对于我来说，研究、健身等事情不是要‘坚持’才能做到的事儿，而是我迫不及待要去做，这些都不是很难受的事儿。

问：国内不少学生都想当您的学生，那您是如何挑选学生的呢？

Eric：首先我要通过你向同学们对我们工作的兴趣表示感谢！我很希望能满足同学们的愿望。有时候，这不仅仅是学生的问题，也是我自己的精力和资金的问题。在美国，招学生是一个相当昂贵的事情，每人每年要资助近十万美元。我的组现在有将近20个学生，再大的话，钱也不够，我的精力也不够。我带学生还是比较认真的，比如每星期至少有一次讨论，20个学生就有20次，一星期排下来也差不多了。在挑选学生的时候，我比较注重文化，风格，兴趣的多样性和平衡，而非单纯的高产，不会偏重某个国家人特别多，也力求性别的平衡。除了这些原因，还需要一些研究兴趣的匹配，以及研究之外的东西。我招学生的时候，可能会注重一个人的个人修养或人品——就是他/她作为一个普通人的价值，而不仅仅是作为研究者所拥有的能力。研究方面的能力是可以培养的，无论是跟着我还是其他老师，学生的能力都一定会有提高。研究的起点不见得非要高，但是对学生的个人性格或修养有要求，我个人比较喜欢学生在我的组里产生正能量。更细致地话，我比较看重这四点：

一、有独立精神以及独立思考的能力。不是说我师兄师姐在做什么选题、最近最热是什么选题，我都去问问然后听从。敢于卓尔不群，坚持自己的观点；但同时也谦虚理性，而非顽固不化。

二、要有一些理想主义。不是做每一件事都有一个精准的投资回报率，要有一些纯粹为了兴趣和爱好而努力、为了尽善尽美而下功夫的追求。享受过程，而不是苛求结果，把结果当成过程的副产品，有了很好、没有也不沮丧，这样才能真正享受过程。

三、诚实而豁达。有一种简单、率真的性格，不能工于心计。

四、知书达理，有良好教养，礼貌，易于也乐于和别人共事。

我相信拥有以上四点的同学可以成为我组里很好的团队成员。

问：每次您来微软亚洲研究院做讲座，总是座无虚席，人特多。您和微软亚洲研究院有什么渊源？

Eric：微软研究院一直以来都和CMU（卡耐基梅隆大学）有众多的交流与合作。对于我个人而言，微软亚洲研究院至今都是国内为数不多的顶级计算机研究机构，培养的学生大都非常优秀。加上我在这里有很多的朋友，他们也经常邀请我过来和同事学生们交流交流，我也都十分乐意，也学到很多东西。

问：最近在线教育（MOOC）比较火，不知道您有没有打算在网上开机器学习相关课程？

Eric：现在MOOC很火，也颇具争议。从普及知识的角度，Coursera做了一件大好事。但作为想对某个领域有深度了解的同学来说，也必须对自己有一个清楚的定位。如果你的终极目标是对某个话题有一定了解，通过MOOC来学习是个还不错的选择。但如果你要以学习的某个方向为职业的话，就需要更多考量。现在的年轻人比较容易从众，可能会出现的一种情况是，从极度自卑到极度膨胀的距离可能会变得非常短。我个人认为，MOOC更重要的一个作用是普及教育，是扫盲用的。然而作为一个博士研究生，你是需要普及知识，还是更高级的知识？我现在还没有加入在线教育，未来可能会，具体视CMU的政策而定。MOOC在普及知识的方面已经做得非常好了，但不见得是我最需要做的事情了。我尝试提供一些更高端、更适合专业研究需要，听众少一点、技术含量更高一点的课，所以我最近把概率图模型这门课录像、并通过CMU放到网上。Daphne Koller教授也开过这门课，在MOOC上有，但我的课教得比较前沿、更难，内容更多些，进度也可能也快些，提供给已经有扎实基础，需要深入研究机器学习的学生。今后我还会公开一门高级机器学习课。

问：机器学习、人工智能，现在火得不行。您觉得是因为它到时候该火起来了，还是炒作的成分居多？

Eric：二者兼有吧。现在的确是到了机器学习的时代——计算机应用的需求超过了原来普通的编程、普通的数据库所能提供的解决方案。数据很大，人肯定处理不过来，数据库不懂得提供一个足够智能的处理服务。机器学习虽然也没能提供解决方法，但它提供了一个大的思路、有意思的思路。有些人可能对机器学习存在一些比较庸俗的想法见解。比如说停留在‘机器学习到底是不是经典意义的人工智能的问题上，贬低一些现在在机器学习、深度学习领域的工作，认为这些不是真正的人工智能，哲学或物理意义上不深刻。还有一种观点认定机器学习，甚至只是深度学习已经可以解决一切的问题，只不过是缺乏一步一步去实现。这两种观点都不够成熟客观。我们需要更冷静一些，平台是对的、大方向是对的，但是还是有很多空白需要去填补、很多突破还没有实现。不要急着去摘一些果子，或者闪开就算过了，其实可以扎进去做一些有深度、有分量的东西。

问：作为华人，在美国做科研，会有民族的包袱吗？

Eric：我个人在文化上是非常东西兼容的，对东西方文化都有认真学习和较广泛的了解和心得。在文化层面，我可能较很多国内朋友更接近中国传统知识份子——对中国传统文化非常了解和欣赏，爱古文，诗词，历史；欣赏秦汉，魏晋之风。但说到学术研究，我不希望它被包装成民族之间的比赛。做科研，不能做成比赛，而是一种对未知和完美的自我追求。当成比赛本质上是功利性的，是做不出来一流成果的，更需要对问题有一种本质性的好奇，希望去理解、解决这

个问题。比别人快、击败别人只是副产品。我老跟我的学生说一个例子，写论文与做研究就像花样滑冰、体操比赛一样，如果老想着我的动作能得10分、满分、对手怎么样，你的动作肯定走形、做不好。如果你想着——我要在舞台上享受表演，你肯定会做得很好。

问：在您的研究领域中，数学重要吗？扮演着怎样的角色？

Eric：重要！数学从方法论上提供解决方法，从思维上来讲它可以培养一种严谨的态度，从而为问题寻找一个比较严格的、普适的、甚至更好的解决方法。但是我不太建议单独强调数学或物理某一个学科知识的作用。就像在战场上，我身边有武器，什么合适我就抓起来用，而不是说我一定要从始到终用某一种武器。在研究的时候，我需要寻求一个经得起考验的解决方案，它是从数学来、从工程来、从物理来都不重要。只是，当你在寻求解决方案时，你会发现数学是一个避不开的因素，它自然而然地影响你的选择。

问：最后，能否给有志于从事计算机科学研究的学生提些您的个人建议？

Eric：其实之前了分享过很多次了，但都是以英文的方式。今天借着这个机会，希望也能让更多的中国学生们看到：

一、清楚地认识到什么是你的目标以及什么是你在技术能力和心智上的优势和不足，这两点对于你的成功同等重要。根据这些分析，再来设定自己的目标，规划自己的行动。

二、当你发觉有必要做出改变时不要犹豫。改变自己的职业方向永远不嫌太晚，只要这个改变是严肃的，且你能够保证在新的方向上全力以赴。那些浪费在犹豫、等待上的时间经常会比你去适应新生活的时间更长。我曾经处于自己新职业方向的谷底，因为在我更换专业的时候，我对机器学习（machine learning）一无所知。但我知道，我会爱上它，因为这个领域的美与力量。从机器学习方向的入门者到成为CMU的教授，我花了五年时间。

三、你需要非常努力地工作。在我的研究生和之后的职业生涯中，一周7天，我每天平均工作12小时，并保持着高度集中的注意力和高效率。（当然，我其实并不把我做的这些看做工作，我觉得它们和娱乐活动同样甚至更为有趣。）

四、你努力工作的动力不应该来自你的上司给予的压力，而是因为你受到工作的启发与触动，并爱上了工作。我经常告诉我的学生和朋友，我的研究就是我的生活，而不是仅仅把它看做工作。相比于看电视和玩纸牌，我更享受从事研究。

五、与此同时，你仍然可以拥有一个多彩的、充满活力的生活。我是体育迷和音乐迷，也喜欢做其他许多事情，享受和家人、朋友在一起的时光。所以我从不消磨时间。当我不在工作时，我尽情地娱乐或锻炼身体（比如去健身房，弹钢琴……），并花时间和家人、朋友们相处。

六、你需要让自己变得非常有创造力，非常独立。实话说，我研究生阶段的很多研究点子并不来自于我的导师，而是我自己想出来的。我从Berkeley的导师那里获得的，更多的是研究上的精神和资金支持、研究方式和研究品味上的启发、诚实的品质和自豪感，以及他们的友谊。

七、要了解你研究领域中最新的进展。每年读一千篇论文并不是一件耸人听闻的事。事实上如果想成为领域内的顶尖学者，你要了解一切，不仅在本领域，还有许多的相关领域。

八、当你做出突破性成果的同时，你也应该尊重领域内权威们的成果。不要认为他们愚蠢而自己更聪明，并且要去问自己为什么多年以前他们没有像你这样做。另一方面，当你确信你确实获得了权威们没有的洞见时，你也要有超越他们的勇气。

九、要变得学识广博而灵活。许多问题和技术都是相关的，不要在一棵树上吊死。

十、在研究中做一个完美主义者或理想主义者从来不是一件坏事。拥有一个乐观的性格和强健的体魄是非常重要的。不要愚蠢地假设自己因为比别人更聪明，所以能不那么努力。事实上，大多数我认识的顶尖科学家都是既绝顶聪明，又会长时间工作并保持高效。所以你需要健康的身体来与他们竞争。我自己在大学时代就是一个半专业的运动员。比起那些比我年轻10 - 20岁以上的学生们，我至今仍能更高强度、更长时间地连续工作和娱乐。

十一、学会沟通的艺术。乐意去与同事和竞争者讨论、分享你的观点。不仅向同事学习，也要向竞争者学习。

十二、最重要的是，请保持诚实、开放、耐心、快乐和富于远见，而不是变得卑鄙、封闭、急躁、苦涩和短视。

人物简介

邢波

邢波（Eric Xing）博士现任卡耐基梅隆大学计算机科学系教授。他的主要研究兴趣集中在机器学习和统计方法论的发展及大规模计算系统和架构，希望能够解决在复杂系统中的高维、多峰和动态的潜在世界中的自动化学习、推理以及决策问题。邢波教授拥有美国罗格斯大学分子生物学博士学位和美国加州大学伯克利分校计算机科学博士学位。

漫谈2014年人机交互（CHI）大会

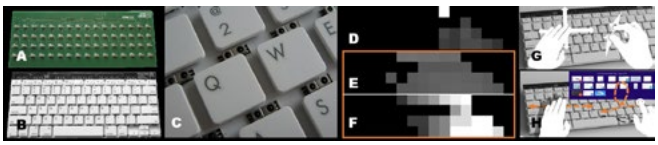
文/Darren Edge

编者按：2014年度以人机交互为主题的顶级会议ACM SIGCHI已经落下帷幕。微软研究院在此次会议的入选论文总数仅次于卡耐基•梅隆大学，位列第二。此次会议中，有哪些创新想法或技术让人眼前一亮?听微软亚洲研究院主管研究员Darren Edge给我们娓娓道来。

如果这篇文章有幸被你阅读到，那么我猜想它十有八九是被发布在了某个网页上。很可能你在用物理鼠标或者触摸屏上下滚动网页来阅读，打开网页需要一款浏览器，而浏览器要运行在操作系统之上，操作系统离不开硬件支持，这些硬件又被包在一个物理“计算机”内（例如个人电脑、平板电脑或移动电话）。但你有思考过这些技术是如何根据用户的知识、技能、需求和价值观等理论而设计的吗？换句话说，你可曾想过，研究人员是如何在实验室和自然环境下理解人们在现实中的使用场景并对其做出响应的？当你在探讨这些问题的时候，你其实已经在思考人机交互（Human-Computer Interaction，HCI）这个课题了，而且我相信正在思考的绝对不只你一人。

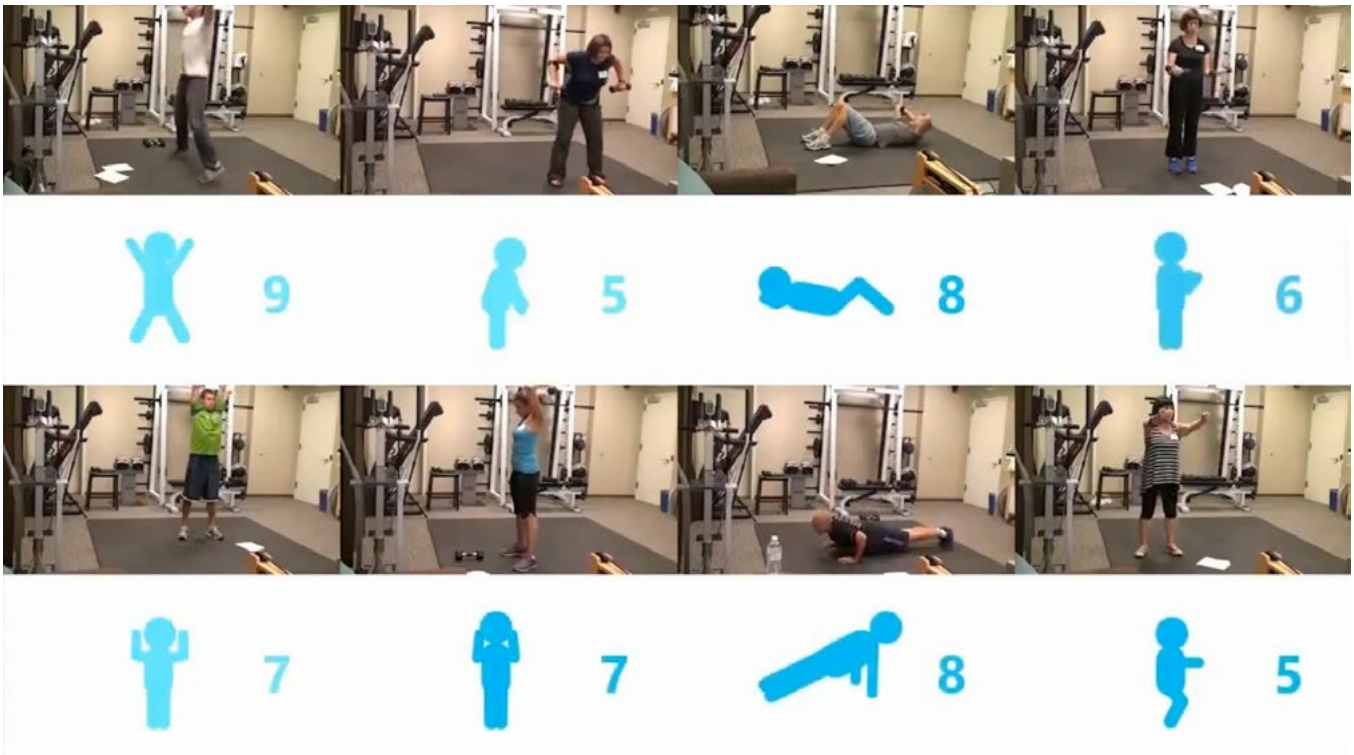
每年，多达3500位研究员会从世界各地聚集在一起，参加一项以人机交互为主题的顶级国际会议——CHI大会（发音类似于“kai”）。今年的CHI 2014在加拿大安大略省的多伦多市举行。为期6天的活动包括2天的研讨会，有多达15个平行活动可选，于是选择参加哪些会议成了困扰参会者的难题。由于不是每个对人机交互有兴趣的人都有机会前往CHI，而且每个参会的人也都只能参与CHI的很小一部分，我希望借此机会和大家分享一下我个人关注到的亮点。

更高效的案头工作



我参与此次CHI大会的首个环节是以“周边互动：塑造研究与设计空间（Peripheral Interaction: Shaping the Research and Design Space）”为主题的研讨会。这次研讨会让我尤其兴奋，因为“周边互动（peripheral interaction）”这个术语正是我在2008年的博士论文《周边互动的有形用户界面（Tangible User Interfaces for Peripheral Interaction）》中首次提出的。在立场文件（position paper）中，我回顾了自己早前对周边互动的定义：用户与其工作空间和注意力范围内的周边对象进行快速而频繁的互动，进而提出了一个用于描述周边互动质量的一般性框架。身处桌面工作空间的情境下，这些周边互动的质量显得更为重要——尽管移动和遍布式计算已经有了长足的进步，我们还要花大量时间在办公桌上使用传统PC和笔记本电脑进行工作。

从工作的角度看，CHI上展示的几个项目试图让桌面交互更加流畅、高效，以提升工作的效率，特别是其中两个项目对如何改善普通键盘的效用提供了思路。第一种是“击键-悬停-滑动（Type-Hover-Swipe）”，这种改良的机械键盘可以识别与按键接触或者按



键上悬停的手势。这个项目由微软剑桥研究院的Stuart Taylor及其同事完成，并且有幸赢得了最佳论文奖（Best Paper Awards）的殊荣，另一个键盘项目被称为GestKeyboard，可以在未改良的键盘上识别轻抚键盘的手势，并且可以与常规打字动作无缝结合。这个项目的第一作者是2011年曾在微软亚洲研究院人机交互组当过实习生的Haimo Zhang。

更全身心的玩耍

案头工作是人机交互研究的传统研究领域，近年来，不少研究者的研究兴趣转向了案头工作以外的互动设计，不局限于工作。其中最令人兴奋的人机交互研究趋势是体育锻炼类游戏（exertion gaming，或简称为exergaming），将体育锻炼、游戏和社交互动融入在一项活动中。

由Dan Morris和微软雷德蒙研究院其他同事合作的Exercise Tracking系统借助可穿戴传感器来寻找、识别和计算重复性练习的次数。另一个项目也颇有意思，它基于运动的系统也成为本届大会的本科生研究项目大赛优胜者，我有幸（也有压力）担任了评委。Kyongwon Seo基于自主的康复设计运用了微软Kinect设备和游戏化元素，鼓励中风患者在自己的家中继续康复。这也是会议期间展示的诸多Kinect创新用法之一。

另外一个令我特别感兴趣的系统则关注更加娱乐化、体态化的互动。这个项目名为VacuumTouch，是作者Taku Hachisu在2013年在微软亚洲研究院实习期间与人机交互组研究员Masaaki Fukumoto合作的成果。他们的系统在一个特定方面比常规的触控表面更具吸力——它采用一只空气泵和一套电磁阀空气阀，借助吸力来移动和限制用户的手指。

更情境化的学习

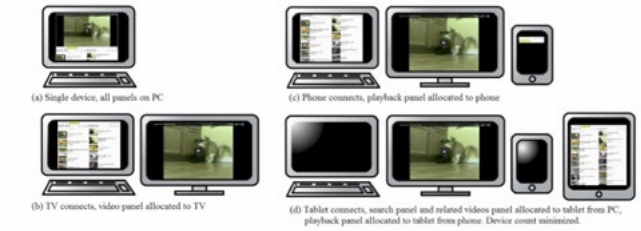
作为一名研究人员，我一直密切关注人机交互及相关领域的技术前沿，希望能够不断磨砺自己在模型、方法、概念、理论和框架等方面的想法和能力。在像CHI这样的会议上，这项工作变得十分容易——因为你需要做的就是聆听并向在场的顶级专家学习。然而，在平常的工作生活中，人们往往比较难以找到时间和动机去学习，因为学习毕竟不是一件易事。在以前的研究中，我就曾试图探索如何利用一天当中短暂而稀疏的空余时间进行“微学习”，帮助人们解决学习第二语言的难题。在CHI 2014期间，我发现了几个在解决现实环境下终身学习问题的、有潜力的项目。

有两篇论文特别关注了情境化词汇学习的挑战。第一个项目是智能字幕（Smart Subtitles），它能提供专为语言学习者设计的交互式视频字幕。第二个项目是WADE，这是一个集成开发环境（IDE），它可以自动修改现有软件应用程序的用户界面，例如将用户界面上的标签和文本翻译成另一种语言。在研究母语和非母语识

别的论文中，有一篇研究了实时多方对话中自动化文稿（Automated Transcripts）分享的效果。这项技术不仅在开会时使用会很方便，也将有助于像我这样的外国人在北京的日常交流，因为我的英式口音往往很难让别人理解。

除了学习特定领域的知识和技能外，学习如何管理时间和精力也是很重要的。有一个项目受到了微软亚洲研究院与韩国科技和未来规划部（MSIP）联合计划的资助，微软亚洲研究院副研究员Koji Yatani与来自KAIST的同事共同探索“大学生如何迷上智能手机”。他们发现，面临成瘾危险的学生每天使用智能手机的平均总量高达111次，合计4个小时。并不是所有参与者都认为自己花在这方面的时间是富有成效或有明确的目的。这使我想起了一个项目，可以用有趣的方式追回一部分在这方面损失的时间，并用于更有意义的活动。这就是正在进行中的个人任务自包研究（Selfsourcing Personal Tasks），该项目由Jaime Teevan和他在微软雷德蒙研究院的同事们合作进行，旨在帮助人们将众包方式运用于他们自身，将大量的个人信息搜寻任务分解成易于管理的“微任务”。就像“微学习”一样，该项目有助于在短期和长期内维持用户的动机和投入。这些方法通过降低有成果和有目的手机互动的门槛，让人们感觉到自己对智能手机使用掌握了更大的控制权，而不是被智能手机所控制。

更吸引人的演示



正如我前面提到的，要想最大限度地利用参加CHI的机会，最严峻的挑战之一在于从众多平行活动中做出取舍。如果你感兴趣的演讲出现在平行会议中，决策过程就会进一步复杂化，意味着你必须制定一个复杂的会议转场计划，但这又会给你和每场会议中的演讲者带来不便。

研究人员一直在努力解决会议环节编排的问题，让各个由密切相关演讲组成的环节在时间上不相互重叠。

现在Lydia Chilton及其合作者已经能够帮助人们简化这一过程——其方法就是利用人群的力量——尤其是参与召开筹备会的议程委员会（PC）成员。她的Frenzy系统在CHI 2014期间赢得了荣誉提名奖（Honorable Mention Award），可用来对论文进行环节分组，并形成最终的CHI 2014会议议程。我在“演示技术”论文环节中，直接体验到了Frenzy的好处，因为该系统帮助PC委员会成员成功地将我的两篇基于PPT演示的论文与我们微软雷德蒙研究院的同事提交的一篇密切相关的论文安排在一起。

这三个相关论文涉及到的都是具有挑战性的活动：对演示的叙事结构加以规划（TurningPoint）；通过结构化的筹备和排练来准备演示（PitchPerfect）；以及为现场观众进行软件演示（DemoWiz）。显然，微软公司内进行着很多与演示相关的工作，并且为PowerPoint这样的产品指明了有趣的新方向。敬请期待！

我要推荐的最后一个项目用充满乐趣、引人入胜的方法介绍了一篇优秀的论文，让所有人都从中受益——事实也是如此，它赢得了“观众评选最佳演说奖（a People's Choice Best Talk Award）”。Panelrama项目提出了一个跨设备Web应用的全新开发模型，并且借助一款演示应用和最新的可穿戴技术进行了演示。这个项目意味着我们对跨设备体验的思考方式又向前迈进了一大步。演示者Jishuo Yang把额外的设备接入他的演示应用时，组件立即动态地重新分配，以确保界面控制板与交互设备之间的最佳契合。在最后的演示中，Jishuo用他的手表读取即时信息和控制幻灯片进度，用头戴式显示器阅读讲稿，用手机查看幻灯片列表，并用连接到投影屏幕的笔记本电脑播放演示文稿的当前页面。总的来说，这个对强大跨设备交互功能的演示给人留下深刻印象，而所用的技术则都是相对简单的HTML扩展。

结语

以上这些就是我在CHI 2014大会期间所观察和学习到的人机交互领域最前沿的技术成果，我们的研究人员关注到了人们工作、生活、娱乐和学习的方方面面，或许这仍然只是窥探到这个充满魅力的研究领域的冰山一隅，更多奇妙的科技还有待富于创新、敢于挑战的研究员们去触碰、去发现。

作者简介

Darren Edge

Darren Edge现任微软亚洲研究院人机交互组主管研究员。Darren的研究主要以设计为导向、注重对重要的人为活动在主观和抽象上分析，并希望能以此开发交互的系统帮助实现这些活动间更好的转换。他的研究兴趣集中在探索技术如何支持跨越不同领域的学习和沟通，如第二语言学习、体育锻炼类游戏、案头工作、和现场演示。Darren于2008加入微软亚洲研究院。在此之前，他就读于剑桥大学，先后获得计算机科学和管理研究学士学位、计算机博士学位。

Microsoft Research
环保心 合家亲
2014 MSRA
FamilyFest

今年，微软亚洲亚洲研究院的家庭日以“环保一家亲”为主题，倡导每个家庭都在生活的小细节中践行环保理念。同时，通过各类有趣的小活动，令小朋友们从小就树立起环保观念。

Matrix释义

Matrix契合微软亚洲研究院的定位——Matrix意为“母体”、“源头”。时至今日，微软亚洲研究院已成长为世界一流、亚洲顶尖的计算技术基础研究机构——它不仅是一流学术成果的孵化器：在世界最权威的学术会议和刊物上发表了上千篇论文；拥有数百项专利；开发的上百种技术被应用于微软产品中，包括微软新一代操作系统、Office、Xbox、必应、Windows Phone等软硬件——还吸引了一批世界级学者加盟，如深蓝项目之父许峰雄等。当前，超过230名中外科研精英正在为它效力。微软亚洲研究院因此被誉为“全球最‘火’的计算机实验室”（《麻省理工学院技术评论》）。微软亚洲研究院——世界级创新科技的母体和基础研究成果的发源地。

我们所从事的是计算机科学基础研究。Matrix是计算机输入、输出间的交叉网。将原《微软亚洲研究院通讯》更新为《Matrix》，亦因微软亚洲研究院汇聚了来自全球各地的科技精英，并将中国和世界最前沿的研究连为一体。我们希望，新的Newsletter也能成为连接朋友和伙伴的一个纽带。

Matrix在数学中的含义为“矩阵”。数学是每一位计算机科研人员所须具备的基础能力。

微软亚洲研究院的大多数员工都钟爱《Matrix》（黑客帝国）系列电影。

地址：北京市海淀区丹棱街5号，微软大厦二号楼 100080
电话：86-10-59178888
传真：86-10-82868527
网址：http://www.msra.cn/
博客：http://blog.sina.com.cn/msra
微博：http://t.sina.com.cn/msra
人人网：http://page.renren.com/600674137
视频：http://u.youku.com/微软亚洲研究院/
欢迎将您的意见和建议反馈给我们，有任何问题请联络研究院公关部