

# 整合会话分析与文本挖掘技术来评价协作学习\*

——访谈卡耐基梅隆大学著名教授卡洛琳·佩恩斯坦·罗泽

□王阿习 王旭



编者按：卡洛琳·佩恩斯坦·罗泽（Carolyn Penstein Rosé）是卡耐基梅隆大学（Carnegie Mellon University）语言技术系与人机交互系的教授，是国际学习科学协会的前任主席和创始会员。目前罗泽教授分别担任国际期刊《计算机支持的协作学习》（*the International Journal of Computer Supported Collaborative Learning*）和《电子和电气工程协会学习技术汇刊》（*the IEEE Transactions on Learning Technologies*）的执行编辑和编辑。罗泽教授的主要研究领域包括文本挖掘与分析、计算社会语言学、计算机支持的协作学习和大规模在线课程的技术支持等。同时，她还承担计算机支持的协作学习、语言分析的计算方法、应用机器学习、文档和大规模交互文本信息的自动提取等课程的教学工作。

**摘要：**建构主义学习理论认为协作发生在学习过程的始终，协作学习过程也是会话的过程。协作学习活动的规范设计与顺利实施有利于促进学习者之间知识共享与协同建构。如何评价协作学习是当前教育技术领域研究的焦点问题。围绕整合技术进行协作学习评价问题，我们访谈了卡耐基梅隆大学著名教授卡洛琳·佩恩斯坦·罗泽。罗泽教授认为会话在协作学习情境中具有独特价值，不仅是一种使思维清晰可见的方式，而且参与者的多样化视角有利于知识的协同创新。罗泽教授的研究聚焦于整合语言学、教育学、心理学等相关理论，深入理解协作学习中会话的社会及其实际本质，搭建提高人与人、人与计算机之间会话效果的计算系统。其研究视角是运用社会语言学和会话分析理论设计语言表征的方式和新的计算模型，使语言模式可以被机器学习；研究内容是从计算的视角分析语言与社会之间的关系，开发用于评价在线协作学习过程的工具（如TagHelper和LightSIDE），以便为协作学习的组织者或促进者提供适合特定情境的干预机制和反馈报告。罗泽教授及其团队的重要贡献还体现在优化协作学习过程的自动化分析方法，促进协作学习支持技术从静态支持向动态支持范式转变，研发新的干预机制与动态支持技术促进大规模协作学习等。

**关键词：**评价协作学习；会话分析；文本挖掘；机器学习；计算社会语言学

**中图分类号：**G434 **文献标识码：**A **文章编号：**1009-5195(2017)06-0003-08 doi:10.3969/j.issn.1009-5195.2017.06.001

\***基金项目：**教育部哲学社会科学研究重大课题“‘互联网+’教育体系研究”（16JZD043）。

**访谈者简介：**王阿习，博士研究生，北京师范大学教育技术学院，北京师范大学未来教育高精尖创新中心（北京 100875）；王旭，博士研究生，美国卡耐基梅隆大学计算机学院人机交互系（美国匹兹堡 15213）。

## 一、采用多种研究视角分析基于会话的协作学习过程

### 1. 会话有利于促进知识的协同建构

访谈者：尊敬的卡洛琳·佩恩斯坦·罗泽教授，非常感谢您接受我们未来教育高精尖创新中心

的访谈！会话或语言交互是您研究中的一个重要概念。在社会和教育实践中鼓励语言交流是您研究工作中一个特别重要的主题。为什么您认为会话或语言交流对协作学习至关重要呢？

卡洛琳·佩恩斯坦·罗泽教授：正如你们所提到的，我的研究聚焦于如何更好地理解会话的社会

及其实际本质，并利用这种理解搭建提高人与人、人与计算机之间会话效果的计算系统。为了达成我的研究目标，我整合和扩展了会话分析的计算方法、文本挖掘技术、会话机器人技术和计算机支持的协作学习的方法。我的研究团队是高度跨学科的，已发表200多篇同行评议的文章和书籍，迄今已在语言技术、学习科学、认知科学、教育技术学和人机交互5个领域颇有建树，而且在其中三个领域发表的论文均已获奖或得到获奖提名。

我的研究目标是开发支持有效参与会话的技术，从而对人类的学习、成长和福祉产生积极的影响。会话在教学情境中特别有价值，部分原因在于它是一种使思维清晰可见的方式，并且会话双方可以协同建构新的想法和创意。通过会话产生的新想法可能来自于参与者的多样化视角，尽管与他们在互动之前每个人所拥有的想法都是不同的。当学生认真对待自己的想法和他人的想法时，他们也会认真倾听，并将自己的思维模式与他人的思维模式进行比较。通过这一过程，他们可能会意识到自己的理解与同伴理解存在差距。当他们通过讨论来寻求弥补这些差距的方法时，就可能产生学习绩效。只有通过讨论，不同学生的思想才能汇聚在一起，做出共同反应并受益于彼此独特的视角。

## 2. 整合语言学和教育学的视角分析基于会话的协作学习过程

访谈者：您刚才提到会话是人际互动中一个复杂而动态的过程。那么您认为语言学和教育学情境中会话分析研究的区别与联系是什么？您是如何整合不同视角来分析基于语言的协作学习过程的？

卡洛琳·佩恩斯坦·罗泽教授：语言是复杂多样的，就像盲人摸象中的大象一样，每一个盲人都依据自身的经验用一种完全不同的方式来描述所感知到的东西。我在会话分析方面的正式训练来自于计算语言学。我从语言学的视角（社会科学的子领域）分析会话研究，并将其与教育科学视域中的会话研究进行对比。

对于协作学习分析的研究涉及多个领域，包括学习科学、教育心理学、组织行为学、社会语言学以及语言学的其他领域、语言技术和机器学习、社会心理学、社会学以及社会科学的其他分支。每个

研究领域都有独特之处，同时也有一系列特定的研究问题和方法论，这些问题和方法促使研究者从不同的视角分析协作学习。即使在教育科学领域，也有多种研究方法必须考虑。当我们思考语言交互是如何促进学习、如何对学习产生贡献时，我们首先需要思考一个外延更广泛的定义——学习是什么，以及学习是如何发生的。对于这两个问题的回答将影响我们对语言作用的认知，因为语言通常被认为是与协作和学习相关的。除此之外，也将影响我们分析与解决问题的方式，如分析需要解决什么问题，使用哪种方式解决问题，以及最终使用何种方式分析会话过程才能确保其具有较高的信度、效度和效率。教育科学中的分析方法是紧密依赖于教育科学中的理论框架的。

语言学也是一个诸多理论支持的领域。然而，如果我们单从语言学视角分析协作学习，那么我们可能将分析结果纳入一种学习理论。如果我们整合语言学和教育学领域的不同视角研究协作学习，那么我们就能够更清晰地了解人与人之间沟通的语言过程。因此，我们整合各种理论视角分析协作学习，旨在整合各种理论以便更加透彻地展现基于语言的协作过程，而非为了判断哪个视角是正确的。

在我的研究和职业事务中，一个永恒的主题是为语言学和教育学等相关领域搭建沟通的桥梁。因此，相比分析不同领域之间的区别，更重要的是，关注各个研究领域之间的工作是如何相互协作的。识别会话结构是我跨领域研究中的重要组成部分，因为会话结构能够预测并揭示影响个体差异的变量，包括动机、目标、个体和团队的学习成果。我在这一领域的理论贡献是从计算视角重新解释社会语言学和话语分析的大量质性框架，特别侧重于在系统功能语言学理论中描述人际动力学的框架。

为这两个学科搭建科学交流的桥梁已成为我专业服务的中心。例如，在我担任学习科学国际协会主席、国际人工智能协会执行委员等职务期间，我牵头组织了一系列的研讨会，旨在创建一个正式的组织结构，以促进这些不同组织之间的沟通、协调和合作。在此基础上，学习科学协会联盟得以成立。此联盟囊括了国际学习科学协会、人工智能教育应用协会、学习分析研究协会、国际教育数据挖掘协会、计算机器

学习协会等。这个联盟汇集了来自世界各地的数千名研究人员，他们分别从以人为本和以技术为中心的视角研究技术促进学习的效果。该联盟近期努力的成果是将在2018年6月的伦敦学习节中举办三个会议，以此作为伦敦学习节的一部分。

## 二、整合会话分析与文本挖掘技术评价协作学习的效果

### 1. 连接会话分析与文本挖掘技术：会话交互自动分析与文本的社会性分析

访谈者：您认为文本挖掘技术与会话分析之间的关系是什么？您是如何整合二者的优势来分析协作学习的？

卡洛琳·佩恩斯坦·罗泽教授：广义上的文本挖掘是一个综合运用从信息检索到文本分析等方法的领域。会话分析及其相关的子领域聚焦于应用理论框架讨论分析数据。我的工作是将这两个领域联系在一起。我研究的核心技术贡献在于会话交互的自动分析和文本的社会性分析（如观点建模、语义分析和意见挖掘）。我把这些问题称为语言的社会性解释。我的研究团队在此方面做的研究工作已经发表在计算语言学和教育科学领域的刊物上。

我的团队有着独特的研究理念：运用社会语言学和会话分析的理论设计语言表征的新方式，这是实现基于语言自动的社会解释的基础。设计反映这些理论观点的计算模型使得语言模式可以被机器学习。我在这个领域的早期工作验证了一个概念：即应用于原始通信数据的机器学习可以复制多维的方法来分析协作过程，这些协作过程在CSCL社区中被认为是具有影响力的。计算语言学社区的研究人员致力于大规模的社会互动分析，而运用机器学习分析协作学习过程使我的工作计算语言学社区中与众不同。2016年我们团队刊发在《计算语言学杂志》上的文章详细阐述了计算社会语言学这一新兴领域的前景，以及我们在研究中整合的、构建的和扩展的理论、方法论与模型建构技术（Dong & Jong, 2016）。另外，我在计算社会语言学领域的研究成果已于近期发表在《自然》杂志上（Rosé, 2017）。

访谈者：您刚才提到计算社会语言学，属于一个融计算机科学、社会学和语言学为一体的交叉学科，

请您简要介绍一下计算社会语言学的核心观点？

卡洛琳·佩恩斯坦·罗泽教授：随着越来越多计算语言学领域的研究者试图从社会学的视角解读语言，社会语言学领域的研究者逐渐意识到使用计算语言学的方法和技术可以改进本领域的模型，提高研究成效。同时，已有研究证明，在计算机模型中考虑语言社会意义的变化，可以完善现有的自然语言处理工具。社会语言学家研究的大量理论和实践，也能够影响计算语言学领域的基本方法。这两个领域的研究边界并不是截然分明的。计算语言学领域日益表现出对分析和建模语言的社会维度的兴趣，鼓励社会语言学领域和计算语言学领域的学者之间开展多样化合作。然而，到目前为止，这两个领域之间的协同作用潜力、共性与互补等都尚未被系统地探索。在这种情况下，一个多学科交叉的新兴学科应运而生，我们把它称为“计算社会语言学”。

计算社会语言学是一门交叉学科，从计算的视角研究语言与社会之间的关系，其研究内容包括开发支持社会语言学家的工具，确立建模与分析语言内容与社会情境信息等相关数据的新统计方法，采用社会语言学理论与方法开发或改进自然语言处理工具等（Dong & Jong, 2016）。

计算社会语言学的关键思想是从社会语言学和话语分析互动的大量理论模型中汲取见解和属性，并使用多种计算方法分析这些见解，以获得能够影响社会学发展的最重要的本质。我的研究方法始终以调查会话的过程、形成合适的理解会话的模型为中心，这些模型足够精确、可复制，并且对真实世界具备较强的解释力；下一步是改写、扩展和应用机器学习和文本挖掘技术构建计算模型，运用计算模型自动分析自然语言交互的特点与规律；最后，随着自动监控自然语言交互技术的发展，逐步建立对真实世界有益的干预措施。

我的研究工作主要是使用计算语言学中广泛采用的大规模数据驱动的方法辅助现有的社会语言学研究，并运用社会语言学改进和挑战计算语言学采用的方法和假设。我们已经证明了计算语言学和社会语言学这两个研究领域之间相互协同的潜力。我们非常希望更多的研究者了解这两个研究领域紧密合作的益处，并对公开的挑战进行讨论。此外，我

们也希望为这个学科的研究者成立一个研究社区，以便保持社会语言学和计算语言学这两个领域之间的联系。

## 2. 采用机器学习工具分析协作学习过程

访谈者：我们从您以前的工作中了解到，您的实验室开发了两个工具：TagHelper 和 LightSIDE，它们是用于分析协作学习过程的机器学习工具。您能简要介绍一下这两个工具吗？

卡洛琳·佩恩斯坦·罗泽教授：为其他研究者提供工具是我的研究团队的长远目标之一。我们研发的大多数工具可以从 DANCE<sup>®</sup>网站上的资源页面获得。我的研究团队最初开发的用于文本挖掘的平台是 TagHelper。该工具能够使用机器学习技术自动分析会话数据。在计算机支持的协作学习和教学交互研究领域，分析学习者的会话过程尤为重要。因此，我们研究的一个重要目标是为研究者提供有价值的工具以支持他们开展这些领域的研究。同样的技术也可以用于评估在线合作学习的过程，以便为小组学习的组织者或促进者提供适合特定情境的干预机制和反馈报告 (Rosé, 2007)。

TagHelper 是我们研发的文本挖掘工具的初始原型。迄今我们已经极大地扩展了 TagHelper 工具的功能，尤其是接口支持的开发过程方面的功能。我们将更新与扩展后的版本称为 LightSIDE，它为一家名为 LightSIDE 实验室的创业公司奠定了基础。过去我们还有一个中文版的 TagHelper，感兴趣的研究者可以联系我们，我们也可以向中国的研究人员提供类似的 LightSIDE 版本，以支持中文数据的文本挖掘。用 LightSIDE 支持开发过程的关键优势是它不仅可以利用机器学习来分析和表征数据，而且还能理解不同方法对数据表示的有效性。此外，我们还设计并开发了使用 LightSIDE 的课程 (Mayfield & Rosé, 2013)，在课程中我们介绍了 LightSIDE 用于判断语言属性质量的指标，以及用于搜索和筛选更有效的语言属性的方法。

## 三、规范设计协作学习活动改善协作学习效果

### 1. 促进协作学习支持技术从静态支持向动态支持范式转变

访谈者：为了改善协作学习效果，您在研究中

采用了哪些干预机制与支持策略，以促进协作学习过程中学习者之间开展积极对话？

卡洛琳·佩恩斯坦·罗泽教授：我认为会话系统想要获得最大的影响力，它必须首先能够理解对话，输出对话，并且融入对话。其次，该系统应当对不同情境下的会话机制有深入的理解，比如在特定情境下，什么样的对话对谈话效果有所增强或削弱。为此，我的研究团队已经对学习科学领域中发展迅速且相互关联的两个研究方向做出重要贡献，这两个研究方向分别是协作学习过程的自动分析和协作学习的动态支持，另外我们还使用智能会话机器人支持情境感知的协作学习。

至少已有 10 年的研究成果（包括我自己的研究）表明，当计算机可以自动提供学习支持时，学生可以从与小组成员的互动中获益，特别是在互动和情境感知技术的支持下。虽然在计算机支持的协作学习领域中，最先进的技术仍然是静态的支持形式，例如结构化的输入界面、提示和分配给学生的脚本角色，但是现在为协作学习提供动态支持的技术也已经公开。我的研究团队在促进协作学习支持技术从静态向动态这一范式转变中扮演重要角色，尤其证明了为协作学习提供基于脚本的动态支持要优于其他静态形式的支持。我们的研究发现，与单独提供材料相比，为学习者提供基于会话机器人的、情境感知的动态支持更能显著提高学生的学习成绩 (Adamson et al., 2014)。

我们还采用机器学习技术实现了自动分析协作学习过程中的语言交互特点，尝试建立了学习者参与协作学习的认知、动机与关系等维度的计算模型 (Howley & Rosé, 2016)。另外，我们使用无监督的动态贝叶斯网络为会话过程中的言语风格适应建模，并将其作为言语交互自动分析的一步。在这项关于言语风格的研究中，我们发现交谈者可能会在互动中改变他们的说话方式，变得越来越相似，或者变得不那么相似。通过将言语风格调整作为一种社交暗示，我们可以更好地确定谈话参与者是否在努力构建一个共同的会话基础 (Gweon et al., 2013)。

### 2. 设计规范的交互活动 (Transactivity) 是有效开展协作学习的 DNA

访谈者：您在研究中提到交互活动是有效开展

协作学习的DNA。您所指的交互活动是什么？交互活动与协同知识建构的区别是什么？您是如何设计交互活动促进学生之间的协作学习的？

卡洛琳·佩恩斯坦·罗泽教授：在协作学习环境中，当参与者详细阐述、拓展、提问或反驳他们的合作伙伴提出的想法时，他们会进行一场交互式讨论，以努力达成共识，达成共同的解决方案。在教育心理学和计算机支持的协作学习领域，交互活动已得到较好的研究。Berkowitz和Gibbs为交互贡献提供了一个广义的分类，并为每个类别列举了具体的示例（如下表所示）（Berkowitz & Gibbs, 1983）。

表 交互的示例

交互	定义
反馈请求	你能理解或者同意我的立场吗？
双解释	这是对我们共同立场的另一种描述。
竞争并存	我认同你的观点，但我也保留我的意见。
拓展	这是你对自己的立场所做的进一步延伸或详细阐述。

交互活动只是协同知识建构中的一部分，协同知识建构是个含义更广的概念。学生参与协同知识建构的诸多过程中，只有一部分是交互性的。

小组学习的研究文献提供了强有力的证据，证明了学生之间合作交互的成功依赖于教师通过讨论促进学生主动交流的能力。教师为学生自主学习创造了机会，同时也为学生产生不同的观点与想法提供了必要的支持。在学生会话中，教师能够依据情境呈现课程内容，学生们认为教师展示课程内容的方式与促进他们实现自己的目标是相关的。教师通过创造一个环境，让学生看到他们参与课程的过程，以促进他们了解自己学习的进程。教师还有机会成为一名导师，为学生提供经验和智慧，并为他们导航以便帮助其走出误区。当学生可以内化教师的引导并将其应用于小组讨论中时，学生的收益是最大的。

在我的研究中，我正寻求通过自动化的方式，如以智能会话机器人的形式，促进小组成员之间的讨论，试图模拟教师在小组学习中的角色。尤其是我的研究团队已经开发了诸多计算机支持的协作学习干预措施，如由会话机器人和计算会话分析技术支持的干预措施，当将这些干预措施应用到当地的Accountable Talk<sup>®</sup>职业发展项目中时可提高学习者的学习成效（其效应值在0.35至1个标准差之间，

等同于提高一个字母成绩（B→A）），并有效提升了参与者所在班级的对话质量。Accountable Talk项目的成功验证了这项技术可应用到类似的具有挑战性的职业发展项目中。

自从我们成功实施Accountable Talk项目之后，我们又在初中、高中、城区和郊区的学校、顶尖的大学和一般的大学，以及社区学院和大规模开放在线课程（MOOC）等实施了一系列成功的课堂研究。我们的实践跨越了多个领域，包括数学、科学、工程和社会科学。我们在上述所有情境中都取得了成功。通过在多样化的情境中开展研究，我们已经开发了一系列智能会话机器人的设计原则。例如，个性化的会话机器人可为学生提供支持，并利于学生之间相互交流。此外，会话机器人也有利于学生控制交互时间。Accountable Talk项目中采用的会话机器人能够成功地激发学生之间的协同交互。学生对机器人的修辞策略比较敏感，比如显示出偏见、显示出对其他观点的开放态度和有针对性的启发，因此我们可以利用这种敏感性使会话机器人更好地服务于学生的学习目标。

我们在一所市区学校九年级的生物课程中实施的一个项目能够较好地体现我们研究的性质与影响力，此项目旨在通过构建协作学习的环境改进科学教学的质量。该项目是与Lauren Resnick教授和学习协会（Institute for Learning, IFL）合作开展的，得到匹兹堡学习科学中心的资助。该项目中学习协会实施的Accountable Talk课堂教学促进实践已被证明能够大幅提高学习者的标准化考试分数，并能保持三年以上。在我们开始合作的时候，这些影响迄今为止没有在其他市区学校的教师中出现过。

#### 四、研发新的干预机制与动态支持技术促进大规模的协作学习

##### 1. 研发新的干预机制支持MOOC中大规模的协作学习

访谈者：我们注意到您的研究团队近期已在大规模开放在线课程（MOOC）方面开展了协作学习的相关研究。社交孤立和高退学率仍然是当前MOOC发展面临的难题。您的研究采用了哪些干预机制促进MOOC情境下的协作学习？

卡洛琳·佩恩斯坦·罗泽教授：阻碍MOOC发挥其巨大潜力（即向大众提供有价值的学习体验）的关键在于，它们不能提供有利于持续参与和学习的社会环境（Rosé et al., 2016）。当学生们使用这些新兴的、刚刚起步的在线学习社区时，这种情况尤为明显。我们的研究旨在为应对这一挑战奠定基础，从案例研究和基于社会互动数据的计算建模开始，以产生新的知识，从而为建立健康的学习社区提供新颖、实时的支持，进而促进学习者高层次的参与和学习（Rosé et al., 2015）。

为增强MOOC平台的功能，我们正在研究新的干预机制，这需要开发新的技术。首要目标是开发一种可以适应新兴社区或亚社区结构的沟通媒介——用于形成和维持支持性关系。我们的最终目标是培育探究性学习社区。在该社区中学生可以互相建立支持性联系，进而协同开展推理活动，并将同伴的思维作为 Accountable Talk 课堂中的资源。我们力求通过提供同步讨论的机会、更有效的帮助、基于团队的工程等机制来实现上述目标。

学习分析中的一个重要研究问题是加快数据循环，从而分析学生的需求和改善学习支持。考虑到社会互动在学习中的重要地位，我的研究团队提出了一套完整的操作体系，包括一个用于从多个平台中采集社会交互数据的基础设施，一种用于分析社会关系对学生学习路径影响的概率序列模型，还有一个支持学生获得积极社会资本的社会推荐系统。

计算分析工作的基础是数据的表示。我们已发表的大部分作品都集中于使用聊天数据或转录面对面讨论中的数据，以分析讨论中的合作情况。这些数据都可以采用一种简单的、统一的、呈序列状结构的文本代码段表示，每个发言者贡献一个文本代码段。然而，当扩展到MOOC中的学习或者其他在线环境（如开源社区）中的学习时，讨论的形式可能会变得更加多样化，因为它们被嵌入到各种平台中。它们甚至可能通过多个单独的数据流同时发生。为此，我们提供了一个开源的数据基础设施，我们称之为“Discourse DB”<sup>③</sup>。它可以将多个数据流转换成一种通用的、集成的数据表示方式。接口层（最外层）的语言表征被进一步转换成“话语”（Discourse），嵌入的话语部分由更小的单位“文本

输出”（Contributions）组成，它们彼此之间通过“关系”（Relations）相连，并且“文本输出”又关联于自动生成的“注释”（Annotations）。这种通用的数据表示方式不仅可以整合来自多种数据流的数据，而且能够应用通用的建模技术。一旦我们感兴趣的数据用一种可概括的方式被表示出来，下一步就是模拟学生的轨迹。这种分析方法便于我们识别干预措施，从而增加对学生轨迹产生积极影响的机会。我们为此创建的资源可以通过DANCE社区获得。

## 2. 设计动态支持技术促进MOOC中大规模的协作学习

访谈者：除上述干预机制之外，您还采用哪些动态支持技术来促进MOOC情境中的大规模协作学习？根据您的MOOC的研究，您会为中国的MOOC设计者与开发者提供哪些建议？

卡洛琳·佩恩斯坦·罗泽教授：关于这两个问题，我可以简要介绍一下我的团队近期所做的主要研究工作，以供中国的MOOC设计者与开发者参考。一项工作是利用MOOC兴起的契机，推动教育软件和应用的大规模传播和部署，将我们开发的技术运用到MOOC协作学习中，以便为大规模协作学习提供动态支持。过去几年，我们通过与开放学习机构（Open Learning Initiative）、edX、戴维森学院（Davidson College）和史密森学会（Smithsonian Institute）进行持续合作，已经制定出体系化的部署措施并产生了积极的影响。在继续这项工作的同时，我们发现许多新的挑战与当前MOOC的教学设计假设有关。如果我们将MOOC作为自主运作的课程，而将合作机会嵌入到一种可选的补充活动中，那么这种教学设计方式就会限制学习者协作讨论的机会。而这种机会能为学习者提供直接的社会支持，并利于学习者加强学习承诺，积极进行决策和学习。未来我们的目标是扩大和深化与在线学习社区领导层的合作，因为我们已经制定了跨越正式学习情境和用于信息共享的非正式学习情境的愿景，这些都是领导层非常关注的。

MOOC的高退学率表明，学生们认为没有从MOOC平台的学习中获得足够的价值，因此花费大量的时间进行MOOC学习是不值得的。我们已经看

到,鼓励MOOC中的同学互相讨论,拥有更多社区情境(如整个班级的讨论论坛)的学习体验,有助于在一定程度上解决这个问题。

最近,我们一直致力于运用基于团队的项目课程设计思路为学习者提供更加持久的社会交互。来自两个团队的MOOC数据分析结果表明,即使在最先进的基于团队的MOOC中,其团队的成功率依然比较低。虽然团队领导者和其他团队成员的行为可以预测团队的结果,但已有证据表明,成功率低的问题在团队开始发挥作用之前就已经出现了。因此,必须改进团队的形成过程,以便建立可能取得成功的团队。我们提出一种基于审议的团队形成程序,以改善选择和启动过程,进而调节与增强团队的讨论交流过程。这意味着需要预先给学生分配前置性任务,让学生自主完成任务,然后发布到公共讨论论坛上,以获得其他同学的反馈。在这种情境下,学生必须选择为少数同伴提供反馈意见。其次,使用自动化过程分析工具评估学生之间的交互贡献。最后,使用约束满足算法将学生分配给团队,从而使得分配给同一个团队的学生之间的平均交互程度在学生群体中达到最大化。

这种划分团队的方法旨在解决两个主要问题。一是解决学生无法确定与哪些同伴开展良好合作的问题。该方法基于对学生之间协作潜力的自动评估而为团队合作提供依据。二是有利于解决MOOC中退学率高的问题。过去基于团队的MOOC面临的一个主要问题是团队中关键成员的流失。在这种范式下,团队形成被推迟,因此被分配到团队的学生是那些已经表现出对课程承诺的人。而且我们发现学生有时会表现出这样一种倾向——一旦分配到某一团队,他们就会过分依赖团队成员的帮助,而不是继续接触更广泛的课程社区。在我们设计的范式下,我们鼓励学习者在团队形成之前更广泛地参与到学习社区中,而且他们确实可以从中获益,因而有利于其后续的团队合作学习。除此之外,我们希望学生有了积极的社区参与体验之后,即使在团队组成后,他们也能继续回来参与这种课程学习。

初步调查结果显示,预先分配需要同伴反馈的前置任务和基于自动检测学生交互情况形成团队,这两种情况均对学生参与MOOC学习有较强的影

响。这种方法已经部署在与史密森学会合作的基于团队的MOOC课程中,而且应用结果与实验室环境中进行的验证研究一致。该试点研究范式的持续发展方向是逐步扩大反馈意见的来源,如引入开放的同侪反馈渠道,而非仅限于个别团队成员内部的反馈。学生除了在整个课程讨论的情境下从更广泛的反馈中获益之外,我们研究的结果也表明,当学生处在公共场合而非私人场合时,他们也会提供更高质量的反馈意见。

访谈者:您刚才所阐述的大规模协作学习的干预机制与动态支持技术,以及前面提及的协作学习分析、评价与规范设计等内容,将会给中国的教育技术研究者提供重要的启示。再次感谢您接受我们未来教育高精尖创新中心的访谈!

注释:

- ① DANCE[EB/OL]. [2017-07-30]. <http://dance.cs.cmu.edu>.
- ② Accountable Talk<sup>®</sup>[EB/OL]. [2017-08-29]. [http://ifl.pitt.edu/index.php/educator\\_resources/accountable\\_talk](http://ifl.pitt.edu/index.php/educator_resources/accountable_talk).
- ③ Discourse DB[EB/OL]. [2017-07-30]. <https://discourse-db.github.io/>.

参考文献:

- [1]Adamson, D., Dyke, G., & Jang, H. et al. (2014). Towards an Agile Approach to Adapting Dynamic Collaboration Support to Student Needs[J]. *International Journal of Artificial Intelligence in Education*, 24(1): 92-124.
- [2]Berkowitz, M. W., & Gibbs, J. C. (1983). Measuring the Developmental Features of Moral Discussion[J]. *Merrill-Palmer Quarterly* (1982-), 29(4): 399-410.
- [3]Dong, N., & Jong, F. D. (2016). Computational Sociolinguistics: A Survey[J]. *Computational Linguistics*, 42(3): 537-593.
- [4]Gweon, G., Jain, M., & McDonough, J. et al. (2013). Measuring Prevalence of Other-Oriented Transactive Contributions Using an Automated Measure of Speech Style Accommodation[J]. *International Journal of Computer-Supported Collaborative Learning*, 8(2): 245-265.
- [5]Howley, I., & Rosé, C. P. (2016). Towards Careful Practices for Automated Linguistic Analysis of Group Learning[J]. *Journal of Learning Analytics*, 3(3): 239-262.
- [6]Mayfield, E., & Rosé, C. P. (2013). LightSIDE: Open

Source Machine Learning for Text Accessible to Non-Experts [A]. Shermis, M. D., & Burstein, J. (eds)(2013). Handbook of Automated Essay Evaluation: Current Applications and New Directions[M]. New York: Routledge: 124-135.

[7]Rosé, C. P. (2007). TagHelper Tools: Tools for Supporting the Analysis of Verbal Data[A]. Proceedings of the 2007 Conference on Artificial Intelligence in Education: Building Technology Rich Learning Contexts That Work[C]. Amsterdam, The Netherlands.

[8]Rosé, C. P. (2017). A Social Spin on Language Analysis[J]. Nature, 545(7653): 166-167.

[9]Rosé, C. P., & Ferschke, O. (2016). Technology Support for Discussion Based Learning: From Computer Supported Collaborative Learning to the Future of Massive Open Online Courses[J]. International Journal of Artificial Intelligence in Education, 26(2): 660-678.

[10]Rosé, C. P., Goldman, P., & Sherer, Z. J. et al. (2015). Supportive Technologies for Group Discussion in MOOCs[J]. Current Issues in Emerging eLearning, 2(1).

收稿日期 2017-09-08 责任编辑 刘选

## Integration of Conversation Analysis and Text Mining for Assessment of Collaborative Learning

——An Interview with the Famous Professor Carolyn Penstein Rosé at Carnegie Mellon University

WANG Axi, WANG Xu

**Abstract:** Constructivist learning theory holds that collaboration occurs in the learning process, and collaborative learning process is the process of conversation. The normative design and smooth implementation of collaborative learning activities are conducive to promoting knowledge sharing and collaborative construction among learners. How to evaluate the effect of collaborative learning is the current focus of research in the field of education technology. We interviewed Carolyn Penstein Rosé, a well-known professor from Carnegie Mellon University, in order to analyze and evaluate the effect of collaborative learning. Professor Rosé considers that conversation has a unique value in collaborative learning situations since it is a way to make the mind clear and visible, and furthermore, the diverse perspectives of participants are conducive to collaborative innovation of knowledge. Her research program is focused on better understanding the social and pragmatic nature of conversation, and using this understanding to build computational systems that can improve the efficacy of conversation between people, and between people and computers. In order to pursue her research goals, she integrates and extends approaches from computational discourse analysis and text mining, conversational agents, and computer supported collaborative learning. Her research perspective is the method of using insights from theories in sociolinguistics and discourse analysis to motivate the design of novel representations of language which enables automated social interpretation of language. Designing computational models that reflect these insights makes the patterns in language learnable. Her research content is from the computational perspective to analyze the relationship between language and the society, and to develop tools for evaluating online collaborative learning processes (such as TagHelper and LightSIDE), in order to provide appropriate intervention mechanism and feedback reports for the organizers or promoters of the collaborative learning. The important contribution of professor Rosé and her group also includes: the automation analysis method of optimizing collaborative learning process, enabling the paradigm shift of collaborative learning support technology from static support to dynamic support; the design and development of new intervention mechanisms and dynamic support technologies for large-scale collaborative learning, and so on.

**Keywords:** Assessment of Collaborative Learning; Conversation Analysis; Text Mining; Learning Through Discussion; Machine Learning; Computational Sociolinguistics