

复杂认知：可视分析系统设计与评估的新理论基础

张小龙

(宾夕法尼亚州立大学信息科学与技术学院, 美国, 16802)
(xuz14@psu.edu)

摘要：当前与可视分析系统设计相关的研究基本遵循人机交互领域中交互系统设计的研究范式，采取了用户需求分析、系统设计、以及系统评估等主要研究手段。然而，这一范式下的多数研究存在一个矛盾：基于简单认知行为（如色彩感知、空间关系识别等）的研究方法与面向复杂分析行为（如推理、问题求解、决策等）的研究目标存在显著错配。这种范式错配影响研究的理论性，特别是系统的内部效度和研究工作的外部效度。针对这一错配问题，本文提出需要突破传统的人机交互理论基础，将复杂认知理论作为新的理论基础。本文分析了当前可视分析系统的设计与评估方法对内部效度和外部效度的制约，阐述了复杂认知理论与可视分析任务的关联性，并探索了如何利用复杂认知领域的问题求解理论来指导相关系统的研究。

关键词：可视分析；可视分析系统；复杂认知；研究方法

中图法分类号: TP391.41 DOI: 10.3724/SP.J.1089.202*.论文编号

Complex Cognition: A New Theoretical Foundation for Designing and Evaluating Visual Analytics Systems

Xiaolong (Luke) Zhang

College of Information Sciences and Technology, Pennsylvania State University, University Park, 16802, USA

Abstract: Current research on visual analytics systems largely follows the research paradigm of interactive system design in the field of Human-Computer Interaction (HCI), and includes key methodologies including design requirement development based on user needs, interactive system design, and system evaluation. However, most studies under this paradigm have a contradiction: there is a significant mismatch between the research methods developed for simple cognitive behaviors (e.g., color perception, the perception of spatial relationship among interactive artifacts) and research goals targeting for complex analytical behaviors (e.g., reasoning, problem-solving, decision-making). This mismatch may hurt the theoretical contributions of research studies, in particularly the internal validity of a designed system and the external validity of design methods. To address this challenge, this paper argues for a need to go beyond traditional HCI theoretical foundations and proposes to adopt complex cognition theories to build new theoretical foundations. Specifically, this paper analyzes how current design and

收稿日期: 20**-**-**; 修回日期: 20**-**-**. 基金项目: 基金名称 1(项目编号 1, 项目编号 2,); 基金名称 2(项目编号 1, 项目编号 2,); 作者 1(出生年—), 性别, 学位, 职称, 若是博士生导师/士生导师请注明, 若是 CCF 会员请注明, 主要研究方向为******(每个研究方向之间用顿号分隔); 作者 2(出生年—), 性别, 学位, 职称, 若是博士生导师/士生导师, 若是 CCF 会员请注明, 若是论文通信作者请注明, 主要研究方向为******(每个研究方向之间用顿号分隔); (需按模板格式列出所有作者的个人简介).

evaluation methods in research on visual analytics systems constrain the internal and external validity of research, discusses the connections between complex cognition theories and visual analytics tasks, and explores how problem-solving theories from complex cognition can guide research on visual analytics systems.

Key words: visual analytics; visual analytics system design; complex cognition; research method

目前在可视分析研究领域,系统设计是一个主流方向。在 IEEE VIS、EuroVis 和 PacifVis 等主要国际会议中,此类研究占有很大的比重。这类工作基本遵循了人机交互领域中交互系统设计的研究范式,其研究过程包含了用户需求分析、系统设计、以及系统评估等主要环节:首先根据用户或潜在用户在特定数据分析领域所面临的挑战制定设计需求,然后通过实现一个可视分析系统来满足这些需求,并通过用户评估验证系统在系统功能和用户体验两方面是否达成预期目标。

然而,这一范式下的多数研究存在一个矛盾:研究方法与目标用户的实际行为之间存在显著错配。一方面,需求分析和评估所依赖的人机交互领域的方法通常针对简单认知行为(如色彩感知、交互部件之间空间关系识别)和确定性交互任务(如从界面中选择特定数据),而可视分析系统支持的推理、问题求解和决策等行为超越了简单认知的范畴,具有任务非确定性、需动态调整分析策略等特征。另一方面,系统评估采用的度量同样多是源自于人机交互领域,往往是针对个别单一交互工具的效率、准确性等可用性指标和用户体验,却忽视了对复杂分析行为的支持效果。这种基于简单认知行的方法与面向复杂认知的研究目标之间的范式错配,会削弱研究课题的理论性,例如影响系统的内部效度(系统的有效性)和研究工作的外部效度(设计方法的广泛性)。

解决这一错配问题需要突破传统人机交互理论和方法,结合分析行为相关的科学理论构建新的可视分析系统设计的理论基础。本文提出将复杂认知理论作为新的理论基础,以解决现有方法的局限性。具体工作包括:首先总结当前可视分析系统的设计和评估方法,分析其对内部效度和外部效度的制约;其次阐述复杂认知理论与可视分析任务的关

联性;最终探索如何利用复杂认知领域的问题求解理论来指导相关系统的研究。

1 可视分析系统设计的方法和手段

可视分析系统的研究通常包含了用户需求分析、系统设计和系统评估这几个主要步骤。需求分析的目的是在了解用户的工作特点和痛点的基础上总结用户需求,并提出相应的设计目标。系统设计则是对该设计目标的具体实施,设计工作往往涉及数据处理算法工具、可视化工具以及交互工具。系统评估则是对所设计的系统原型进行评测,以检验该系统是否能达到了设计要求。

1.1 需求分析

需求分析为系统的设计提供了目标和依据。需求分析的方法基本上可以分为两类:基于文献,或基于实际调研。前者可运用于某些常见的、普遍性的分析问题(如对社交网络数据分析、或时空数据分析等)。对于这类问题,文献通常会很充分,通过系统的文献调研,基本上可以确定用户的需求。基于实际调研的需求分析针对的往往是比较新的、还缺乏足够研究的用户行为。本节侧重于基于调研的需求分析工作:问卷、用户访谈、以及实地观察。

1.1.1 问 卷

问卷方法通过将一个统一的问卷分发给被试人群,来获得目标用户的行为数据。一个问卷往往包括多个问题,问卷数据的统计分析结果可以揭示用户群体的一般性行为特性,如在执行某类任务时的代表性行为、对某些工具的代表性意见等等。在一个利用机器学习来支持体育赛事录像事件信息标注的可视分析系统研究中^[1], 用户需求的研究工作就采

用问卷的方式来了解球迷们在观看网球、乒乓球、羽毛球等赛事时所感兴趣的事件, 问卷被分发到两个不同的网络社区的约 600 名成员, 其结果帮助确定了机器学习算法需要关注的视频信息特征。

问卷方法的优点是其结果的优异外部效度。如果一个问卷的被试群有足够的样本、且有广泛代表性, 其结果可以被应用到其他类似的情况。因此一个问卷应该尽可能地发放到不同的社区, 涉及尽可能多的用户群体。

1.1.2 用户访谈

用户访谈是通过与用户对话的方式来搜集需要的信息。访谈的侧重点是让用户从其自身的角度来提供研究需要的信息, 例如如何执行一个任务、如何使用一个系统等等。访谈的对象往往是多个代表性的用户, 访谈的过程可以是完全结构化的或半结构化的。在结构化访谈中, 所问的问题和问题顺序是提前设定的, 不能更改; 而半结构化访谈只预先确定部分关键问题, 研究人员可以控制问题的顺序, 并根据访谈对象的回答来临场提出新的问题。

例如在前述的体育赛事录像标注系统的研究中, 需求分析也采用了访谈方式, 以了解专家用户(如体育研究人员、体育数据分析人员)在分析赛事录像时所关心的信息、所碰到的困难等。

访谈方法的优点是能通过直接与用户沟通来获取准确的、深入的、来自于用户的主观信息, 但访谈结果的成功与否依赖于多个因素, 如访谈问题的设计、代表性用户的参与、访谈过程的控制等。

1.1.3 实地观察

实地观察方法的实施往往在用户的实际工作场地进行, 目的是观察用户执行实际任务时展现出的特点和碰到的问题。问卷和访谈的主要目的是获取和用户相关的内在行为信息, 但实地观察的侧重点在于可能影响用户行为的外部环境因素(如工作场所的工况、用户与同事的交流互动等)。

实地观察所针对的用户和任务也需要就有代表性, 而研究人员在观察过程中可以有不同程度的

介入: 作为一个完全的旁观者而不做任何介入; 作为一个旁观者但会对用户的某些行为作出提问; 甚至直接参与到执行该任务中。不同的做法各有优缺点。作为旁观者, 研究人员可以在观察过程中不影响用户的实际操作和工作效率; 但作为参与者, 研究人员可以通过亲身经历更深入地了解用户在操作过程中的体验和需求。

例如在一个关于对汽车自动驾驶系统的可视分析系统研究中^[2], 需求分析就借助了对专家用户工作的实地观察, 以了解他们在评估自动驾驶系统时所关心的模块、所依据的参数、所碰到的痛点等信息, 而这些信息帮助确立了系统的设计需求和目标。

实地观察方法的优势在于能捕捉真实场景下的用户行为数据。当用户在自然工作环境中执行常规任务时, 其行为往往具有较高的可靠性。然而实地观察过程中可能会出现非目标行为的干扰, 从而引入数据噪声, 增加了数据分析的难度。

1.2 系统设计

系统设计一般是以需求分析所给出的结果为设计目标, 在数据处理工具、可视化工具设计、交互工具等方面的技术开发和实现。

1.2.1 数据处理工具设计

数据处理工具设计的主要目的是开发针对某个设计目标的算法或方法。可视分析系统所面对的原始数据往往具有数据量大、结构松散甚至缺失、固有数据结构与任务所需数据结构不匹配等特点, 难以按照任务的要求直接被可视化。因此, 需要借助数据处理算法或方法来对原始数据进行预处理(如提取有用数据、对数据进行结构化等), 以方便后续的可视化设计。

很多情况下需要的算法已经存在, 设计工作的侧重点是针对已有的数据来实现该算法。如在一个支持对慕课学生网上交流模式的可视分析系统研究中^[3], 主要分析任务需要从学生的交流文本中提取相关的交流话题, 而关于话题提取的算法在数据

挖掘领域已有很多成熟的工作,所以这里所需工作就是选用和实现一个合适的算法,并将其融合到系统中。

但有些情况下也需要去根据任务需求去开发新的算法或方法。这类工作在可视分析领域具有一定的创新性,因此很多可视分析的研究会对相关的新算法给与详细的描述。如在一个关于针对数据中概念漂移的可视分析研究中^[4],数据处理工具方面的工作包括了一个用来检测概念漂移的新方法的设计和实现。

1.2.2 可视化设计

这里的可视化设计指的是可视分析系统中的各个视图的设计。一个可视分析系统往往包含多个视图,而每个视图的设计在本质上与传统的信息可视化设计一致,都是针对某个给定的数据结构类型、某些明确的任务,来设计一个特定的视图。如在前述的慕课交流数据可视分析系统中,有一个用于展示学生沟通关系的特性的网络视图,来帮助用户从各个层次来观察这种关系。该网络可视化工具就是基于传统力导向图的方式实现的。

但因为一个可视分析系统有多个视图,所以研究文献只能对这类视图的设计过程的做简单描述,而无法像信息可视化设计研究文献那样对设计过程进行详细的描述,诸如不同设计方案的比较和选择等细节一般就被忽略了。例如在一个关于社交网络的可视分析系统中^[5],系统有一个新颖的“树网图(TreeNetViz)”,用于支持对一个融合了树结构和网络结构的新数据结构的分析。但该视图只是整个分析系统中的一个视图,所以文章对该视图的介绍侧重于其功能在系统中的作用,对该视图设计的具体设计工作(如设计思想、相关算法等)只是给出了相应的可视化设计工作文献^[6]。

有些研究文献确实包含了新颖的视图设计^[7],但由于侧重点在于可视分析而不是可视化设计,没有对设计的有效性进行评估。

1.2.3 交互设计

系统交互设计包括用户与视图中的某些视觉设计元素的交互,涉及到多个视图的、基于任务的交互,以及支持分析流程的交互设计。有某个视觉设计元素的交互往往局限于单个视图中,多数情况下属于^[8]的基本交互任务的范畴。而支持这些具有普遍性任务的交互可以利用开发工具包的固有功能和接口来实现。例如 d3.js 的诸多内在工具可以很容易地实现数据过滤、加亮等功能。

针对特定分析行为的交互设计要求对分析任务本身有深入的了解,这样才能在系统的整体界面布局、视图之间的同步等设计方面做到优化。例如如果了解针对的用户在执行个任务时常用的算法和视图,界面的整体交互设计就可以将所涉及到算法和视图同时展示出来,而不是让用户去自己去配置。同时,将这些相关的视图设计为多个可协调的视图(multiple coordinated views),使得用户可以快速而有效地获取和任务分析的相关信息。同样视图之间需要协调配合,以支持复杂的分析任务。例如可视分析系统中往往要求多个视图之间的协同,当用户在某个视图中选择了某些数据点时,这些数据点在其他视图中应该与其他数据有所区分,使得用户可以方便地从多个视图中观察同一组数据的特性。这类设计需要考虑任务本身的特点。

交互设计也考虑对分析流程的支持,其目的是全面地支持用户的分析过程中的各个任务,以及它们之间的衔接。目前,这方面工作很多依赖于诸如意义构建(sensemaking)等分析模型。这些模型提出了可视分析中的主要阶段性任务,以及它们之间的关系(如任务之间的呈递、迭代关系等)。根据这些模型,可视分析系统的设计可以考虑相应的工具来支持各个任务,以及任务之间的过渡。例如Pirolli 和 Card 的意义构建模型^[9]中包括了多个相互链接的模块,对于其中的数据收纳(shoebox)、证据文件(evidence file)和关系模式(schema)三个模块,在设计中可以提供相应的数据搜集工具,以便于用户为三个模块独立地搜集相关数据。同时该模型也指出了三个模块之间的联系,如快速浏览数据收纳模块中的数据来为证据模块提供证据、通

过结构化证据模块中的数据来协助用户获得重要的关系模式, 交互设计可以考虑相应的工具来支持这种分析步骤之间的转化和递进。

1.3 系统评估

一个可视分析系统往往通过评估来检验设计是否能达到了设计目标。目前常见的评估方法有案例分析、用户访谈、和实验评估。我们分析了从 2007 到 2021 年间在三个主要可视化领域会议 (IEEE VIS, EuroVis, PacificVis) 所发表的可视分析系统研究的文章, 在总共 230 篇文章中, 发现其中约 48% 的文章在评估工作中使用了案例分析, 大约 32% 用到了用户访谈, 但只有不到 7% 的研究采用了实验评估方法。2021 年之后的文章数据还没有分析完毕, 但初步结果表明最近几年的文章表现出与 2007 至 2021 年间文章类似的模式。

1.3.1 案例分析

案例分析是行为科学研究中心的一种常见研究方法^[10], 着眼于通过一个典型性的案例来探索一个新现象, 或描述一个复杂现象。但案例分析方法不能验证因果关系, 如一个工具或设计对用户行为的影响。

可视分析系统研究中的案例分析评估方法一般是通过描述用户使用所设计系统进行分析工作的实际情况, 来展示系统的功用。往往侧重于描述用户在使用一个可视分析系统时用到了哪些工具、采取了哪些步骤、遇到了哪些问题等等。通常案例分析的描述会借助用户的任务完成情况、以及对系统的满意度等指标来证明系统的效果。例如在一个针对微博信息的、基于地图模式的可视分析系统研究^[4], 评估工作使用了两个案例分析 (一个社会事件、一个政治事件), 详细描述了该系统是如何支持通过微博信息来分析这些事件的演变和发展的, 展示了系统是如何帮助用户快速确定事件的关键人物、地点、时间等重要因素, 并通过用户的信息反馈来支持系统的可用性和有效性。

1.3.2 访谈

访谈是通过对用户的直接沟通, 来获取用户对所设计系统的主观评价。在可视分析研究中, 访谈往往和案例结合在一起, 在用户使用、体验了所设计的系统后, 对用户进行访谈, 深入了解他们对系统的体验和反馈, 为系统对他们工作的影响和作用提供详实的佐证。例如在前述的针对汽车自动驾驶系统的可视分析系统研究中^[7], 评估工作就包括了与专家用户的访谈环节。该环节是在用户使用了设计的系统工具之后进行的, 用户对系统的评价涉及系统所提供工具和分析流程的正确性和直观性、系统中各视图的可理解性等方面。

1.3.3 实验评估

在交互设计中, 实验评估方法本质上是评估一个设计的有效性。在可视化和可视分析领域, 该方法是通过比较用户在使用不同工具或系统执行同一个任务的绩效来比较工具或系统的优劣。由于实验评估方法的目的是检验所使用的工具或系统与改善任务绩效的因果关系, 在实验的设计和实施上要求对可能干扰这种因果效应的其他因素进行严格控制, 如不同组的被试之间的差别、在不同条件下执行任务的一贯性等。但这种对其他因素的严格控制也影响了实验方法的应用范围。

对于可视分析系统的评估, 实验方法的最大挑战在于实验任务的复杂度。如果评估的对象只是一个针对简单任务的视图, 实验设计相对容易。例如当用实验方法来检验相邻矩阵视图在帮助用户理解一个社区的网络特性的有效性时, 实验设计可以对比用户在使用相邻矩阵设计和传统节点链接设计时的绩效, 实验任务可以是诸如寻找一个社区内的子社区、或者确定社区内两个节点的关系等典型网络分析任务。这种实验任务明确, 用户与视图的交互操作水平也大致相同, 实验的设计和过程比较容易。

但当评估的对象是针对一个复杂任务的可视分析系统时, 实验设计就会很困难。例如当评估一个面向人工智能模型的可视分析系统时, 典型任务可能需要用户去从不同的角度去比较若干个模型

的特点，而这类任务的目标可能会模糊（需要比较哪些特征、比较几个模型等），用户选择的任务路径可能也会有差异（有的用户是根据以模型为分析单元，而有的用户是基于数据特征）。如果这些差异性得不到有效的控制，实验结果的有效性就可能会被质疑。而严格控制这些因素会对极大地增加实验的复杂度，甚至会导致实验无法进行。如果为了减少任务路径差异而要求用户必须按照某个特定的分析方法，可能会迫使有些用户采用不熟悉的方法，而无法完成任务。可能正是因为实验评估方法的难度，可视分析系统研究中使用实验评估方法的比例较低。

评估一个可视分析系统时，一个可行的折中方案是对系统核心工具进行评估，而不是测试整体功能。这种方法的优势在于显著降低了实验复杂程度。以前述的体育赛事可视分析系统为例：鉴于完整分析任务的实验评估难以实施，实验评估工作选取了两个核心子任务（基本信息提取分析工具和高阶语义信息整合分析工具），并针对它们分别设计了验证实验。

2 可视分析系统设计方法与研究目的的错配

当前可视分析采用的主流方法面临着一些挑战，其根本原因在于人机交互领域的研究范式与可视分析系统的根本目标存在着错配。本节首先通过对比一般交互行为和分析行为的特征，解释分析行为与一般交互行为的本质差异；然后深入系统设计的主要环节，来分析现有可视分析系统研究方法中存在的局限性。

2.1 简单交互行为与复杂分析行为

人机交互领域的研究方法论主要建立在简单认知行为的理论之上。这类简单认知行为通常涉及感知、注意力和直觉等快速反应机制，典型的表现包括颜色/图形的感知、菜单/图标的选择以及快捷键的使用等基础操作。该领域的一个基本假设是用户已经明确知晓待执行的系统命令，而执行该命令需要通

过计算机系统提供的某种工具或媒介。这个假设在诺曼的交互鸿沟理论中得到充分的体现：该理论提出的“执行鸿沟（the gulf of execution）”特指因为交互工具缺失导致用户无法完成一个目标任务的障碍，而交互设计的一个目标正是消除此类障碍。

然而，可视分析系统所支持的行为本质上不是这类简单行为，通常需要推理、决策、学习和问题求解等高阶认知活动。这类高阶认知活动具有显著的非确定性特征：用户需要不断进行任务界定与策略调整。值得注意的是，支持复杂分析行为的工具并非简单交互功能的叠加。以 Tableau 和 ArcGIS 为例，尽管这些商业分析工在基础交互层面（如改变某个视图元素的渲染颜色，拖动某个视觉部件，或者选择并加亮部分数据点等）有着非常友好的界面设计，但可以熟练使用这些工具并不能确保能顺利完成一个可视分析任务。在以往的 IEEE VAST 的竞赛中，有诸多参赛团队使用了这种商业分析工具，但其分析成果质量却存在显著的差异。

类似现象也存在于其他涉及高级认知行为的交互系统。例如，广受欢迎的 Photoshop 有着非常好的交互设计，也融合了很多用于图像处理的工具。但即使熟练使用该工具，用户仍可能难以完成诸如高精度毛发抠图或电影级渲染等专业设计任务。究其本质，这类工具的交互设计普遍着眼于通用的、一般性的、和简单认知相关的交互任务，未能充分考虑诸如分析、设计等涉及高级认知行为的任务需求。

可视分析行为与一般的简单交互行为不同。可视分析行为往往具有目标模糊、分析策略和路径不确定以及任务迭代等特性，执行分析任务的人员需要不断地反省其行为来审视其分析过程、选择相应的策略、根据中间结果来不断调整策略。显然这种行为的认知模式不同于人机交互中典型交互行为的认知模式（如快速找到一个图标或菜单项等）。

Newell^[11]提出的行为时间尺度理论将人类活动划分为不同的频带（如表 1 所示），涵盖从毫秒级的生物神经活动（生物频带）到跨越数年甚至更长时间的社会变化（社会频带）。其中，人机交互

设计所关注的是认知频带(时间尺度为 0.1 – 10 秒)行为特征^[12], 包括基本的感知--动作环路、以及即时决策过程。而可视分析系统所支持的推理、决策、问题求解以及学习等高级认知活动则属于时间尺度更长的理性频带(以分钟为单位, 可延伸至数小时)。

表 1 人类行为的时间尺度(源于 Newell^[11])

时间尺度(秒)	频带
$10^5 \sim 10^7$	社会频带(social band)
$10^2 \sim 10^4$	理性频带(rational band)
$10^{-1} \sim 10^1$	认知频带(cognitive band)
$10^{-4} \sim 10^{-2}$	生物频带(biological band)

Newell 和 Card^[13]指出当这两类行为所调用的认知机制不同时, 其理论基础也存在显著区别: 传统的人机交互行为主要遵循认知心理学的基本原理, 而复杂的分析行为则归属于有限理性(bounded rationality)范畴。

一般来说, 人机交互设计的核心目标是尽量减少用户在完成交互任务时所需的认知资源。例如, 当用户通过命令行与系统交互时, 往往需要依赖长期记忆来构建命令格式和句法; 而图形用户界面则提供了菜单和图标等交互元素, 使用户可以直接识别相关菜单或图标选去完成命令。这种识别行为需要的认知资源远低于调用长期记忆所需的认知资源。

一些交互设计还力图将时间常数长的任务转化为时间常数短的任务。例如, 在图形用户界面中, 一些基于动画的设计(如动态地展开一个文件夹)在认知层面的作用在于利用动画中的物体恒常性(object constancy), 帮助用户直观理解相关物体的变化轨迹例如某个物体突然出现或位置发生改变)^[12], 而无需额外调动认知资源去解释这种变化, 特别是当变化并非由用户主动驱动时。

从认知理论角度来看, 人机交互设计的目标是充分利用人类思维擅长利用经验和直觉的特性, 以设计简单易用的交互工具。认知心理学研究表明

^[14], 人在思维过程会使用两种不同的认知系统: 系统 1(System 1)和系统 2(System 2)。系统 1 运行速度快, 具有一定的自主性, 其决策过程主要基于过往经验。我们通常说的直觉或下意识的行为, 正是系统 1 的作用。然而, 由于系统 1 的判断依赖经验和直觉, 其决策可能会出现错误, 尤其是当遇到的问题与过往并不完全一致, 仅在表面上相似而在本质上不同时。

相比之下, 系统 2 需要大脑的有意识介入, 并调动长期记忆等额外的认知资源, 以执行推理、问题求解和决策等复杂任务。因此, 依赖系统 2 进行的任务通常正确率更高, 但处理速度较慢。

在人机交互设计中, 目标是尽可能地让用户主要依赖系统 1 进行交互, 以减少认知负担, 使用户可以将更多的认知资源用于更为“重要”的任务。而分析行为通常涉及这类更为“重要”的任务, 因为分析任务往往无法仅凭直觉完成, 需要调动系统 2 进行深入的推理、问题求解和决策。

鉴于典型人机交互任务和可视分析任务在本质上的不同, 我们应该重新审视当前可视分析研究的研究方法, 深入了解这些源自于人机交互的方法在指导可视分析系统设计时可能存在的据线性, 以确保设计更符合分析行为的认知过程的需求。

2.2 当前需求分析方面的局限性

当前在可视分析系统在需求分析方面存在一个关键问题: 缺乏对复杂分析任务的深入了解和解析。基于对前述的 230 篇可视分析系统研究文献, 我们提取出 1048 项需求分析条目, 并将其按针对的交互行为(简单交互行为或复杂分析行为)进行了分类。数据显示, 约 56% 的需求条目聚焦于简单交互行为(如数据筛选、视图控制、用户偏好设置等)。

然而, 仅满足这类基础交互需求并足以确保系统对分析任务的有效支持。这些需求只是分析过程中的操作基础, 但分析效能的核心在于更高级的认知活动, 如分析策略的动态选择和回溯、推理过程的控制等。因此, 可视分析系统的需求分析应当

将重点放在分析行为的复杂任务需求层面,而不是与数据、视图等交互元素相关的基础交互任务层面。

2.3 当前系统设计方面的局限性

系统设计方面的主要挑战在于如何界定这方面工作的创新性。一个可视分析系统的研究需要在系统设计上有所创新,而系统设计通常包括了数据算法设计、视图设计、交互设计、系统整合等方面的设计工作,因此系统设计的创新性往往通过这几个方面的具体工作来得到体现。可视分析系统研究的相关文献中一般会明确地阐述研究的贡献点,通过这些贡献点,我们可以提取一个研究项目在系统设计方面的创新点。

通过对 230 篇相关文献的分析,我们发现文献在阐述研究的贡献时,很少会将视图设计、交互设计、和系统整合方面的工作列为创新点,原因可能在于研究人员都清楚视图设计方面的创新工作属于信息可视化范畴,而交互设计方面的创新工作属于人机交互范畴,所以这两方面的工作都不适合于在可视分析的范围内讨论和评价。而系统整合方面的工作多属于工程性问题,不属于研究的范畴。

但很多文献将算法设计作为创新点。进一步分析表明相关算法方面的工作包括以下两大类:

- 基于一个已有的算法来支持和目标分析任务。这类算法可以是不做改动地直接移植一个算法,例如在支持文本信息分析的系统中用到的话题模型[3],也可以是在某个算法的基础上做适当修改以满足分析任务的要求;

- 创建新的算法。

创建新算法的工作虽然创新性最高,但算法本身的创新性超越了可视分析领域可以评判的范畴。对于实现了某个已有算法的工作,其创新性需要深入的讨论。原封不同的实现一个算法显然不具有创新性;而对已有算法做了改进的工作,如果改进只是适用于所针对的分析任务和数据,难以适用于其他的任务和数据,那么其创新性也很有限。但如果算法的改进具有很强的推广性,那么这种工作类似于创建新算法,其研究贡献应该由和算法相关

的领域进行讨论和评判。

视图设计和交互设计也面临类似的挑战。对于视图设计,如果所做工作只是实现已有的视图,那么这方面工作的不具备创新性。如果所提出的设计是针对某个任务的全新设计,具有一定意义的推广性,那么其贡献更适合于在信息可视化领域内讨论,而不是在可视分析领域。例如,前述的针对社交网络社会性的可视分析系统设计[5]用到了一个树图可视化(TreeNetViz)部件作为其的核心视图部件。而树图可视化针对的是社交网络分析中的一种综合了树和图的特殊数据结构,该视图在可视分析系统中的使用是在它的有效性经过信息可视化研究领域得到检验后[6],才融合在系统中的。

对于交互设计,如果工作只是局限于对用户一般交互过程的支持(如视图之间的协同、视图的排布、全局和局部信息的平衡等),因为相关任务已经被大量研究,其创新性也有限。但如果交互设计的工作是有关某一类分析行为的、具有代表性的新发现(如人工智能算法分析过程中用户与算法之间的互动),那么对该工作创新性的判定不仅需要对分析过程的准确掌握,也需要对相关的认知行为和数据算法的深刻了解,其意义往往超越可视分析领域,也会影响到人机交互、人工智能等领域。

所以对于系统设计工作来说,如何准确地阐述算法设计、视图设计和交互设计对可视分析领域的贡献是一个挑战。如果对它们做单独讨论,我们面临的一个问题是可视分析领域是否有足够的专业能力来对它们的创新性作出准确地判断。而这三者对可视分析的重要性不言而喻,因此对这方面工作的创新性应该考虑三者之间的关系,这就需要一个能把三者融合起来的指导性框架。

2.4 当前系统评估方面的局限性

面向系统设计的研究需要经过严格的评估来证明所设计的系统可以有效地帮助用户完成分析任务。但我们对 230 篇文献分析表明,这方面的工作往往也是研究中最薄弱的环节,因为许多评估工作严格地来讲只是通过对系统功效的描述和借助

用户访谈的反馈来佐证系统的可用性, 而不是采取可信的评估手段去证明系统对用户分析工作成效的作用。究其原因, 是因为评估工作选用了不适当的方法, 而这种方法选择上的局限性只会削弱研究的贡献。

前文提到我们对文献的分析表明当前评估手段主要采用案例分析和访谈两种方法, 而较少运用用户实验, 尽管后者验证系统对提升分析行为效果之因果关系的唯一可靠方法。这一现象背后存在着深层方法论的挑战: 在人机交互领域和信息可视化领域, 用户实验虽然被广泛用于评估特定设计(如界面设计、视图设计、交互方式)对离散交互任务(如洞察系统状态、理解数据特性、发现事件关系等)的改进提升效果, 但面对可视分析系统的多部件、多视图、多模态交互特性时, 传统的实验方法面临多重困境:

- 任务复杂性: 分析任务通常需要组合使用多种可视化工具和数据处理技术;
- 实验控制: 分析过程的不确定性与传统实验设计的可控性要求存在根本矛盾;
- 评估颗粒度: 现有度量指标(如系统可用性量表 SUS^[15])仅能评估整体可用性和用户接受度, 难以捕捉分析行为层面的效能提升(如推理路径的发现、决策标准的优化、用户对分析任务完成度和完成质量的认可等)。

这些挑战并非可视分析系统独有, 而是所有包含多种工具和手段的交互系统面临的方法论瓶颈^[16]。

当前可视分析研究普遍采用案例分析方法的现象, 很可能源于实验方法实施面临的困难。然而现有的案例分析方法多聚焦于系统功能描述, 详细介绍用户是如何组合各类数据处理和视图工具完成特定分析任务。这种案例分析方法由于缺乏对用户任务和使用工具的严格控制, 无法确立系统设计于分析效能提升之间的因果性, 同时这对特定系统和特定任务的个案研究结论也难以推广到一般性设计原则。这种局限性本质上是案例研究方法本身在效度(validity)和信度(reliability)上的固有缺

陷^[17], ^[18]在可视分析领域的具体体现。

访谈方法本质上属于启发性评估(heuristics evaluation)范畴^[19], 这类方法在评估系统特定维度(如界面可用性、任务执行流程等)方面有其特定的价值。在可视分析研究中, 访谈常和案例分析结合在一起, 通过对系统使用者的深度访谈获取用户主观使用体验(如工具可学习性)、系统改进建议(如功能的缺点或缺失)等方面的信息。尽管这些质性信息丰富了评估维度, 但访谈方法与案例方法存在着共同的方法论局限: 由于缺乏对变量的有效控制而无法建立因果推论, 局部个案的描述性数据影响了研究结果的可靠性和可扩展性。

综上所述, 简单交互行为与复杂分析行为在认知本质上的根本差异, 导致传统人机交互研究方法在应用于可视分析时存在系统性局限, 这一局限贯穿需求分析、系统设计和系统效能评估等多个关键研究环节。要突破这一方法论困境, 需要首先深入解析分析行为的认知特性和过程规律, 然后基于相关理论构建针对复杂分析行为的理论基础, 来指导和发展有效的需求分析、系统设计和系统评估。

3 可视分析行为的特点

我们可以从现有对可视分析的定义出发, 来了解可视分析行为的主要特点。当前学术界对可视分析的概念主要基于美国和欧盟的两个纲领性文件。美国方面的定义源自于《照亮前进之路: 可视分析的研究和发展议程(Illuminating the Path: The R&D Agenda for Visual Analytics)》一书^[20], 该书把可视分析定义为针对分析推理并借助于交互可视化界面的一门学科。但这个定义过于简单: 可视就是交互可视化, 分析就是分析推理。欧盟在《把握信息时代: 利用可视分析解决问题(Mastering the Information Age: Solving Problems with Visual Analytics)》的白皮书中^[21], 则给出更为明确的定义: 可视分析通过融合自动化数据分析和交互可视化技术, 支持用户对海量复杂数据的理解, 以及在此基础上的推理和决策。

这两个定义共同揭示了可视分析的核心行为包括理解、推理、和决策。在认知心理学上，这类行为所涉及的认知过程被称为复杂认知^[22]。

3.1 可视分析中的复杂认知行为

在复杂认知领域，学习、推理、问题求解（problem-solving）和决策等行为被公认为典型的复杂认知活动^{[23],[24]}。这种认识与欧盟白皮书对可视分析的定义高度吻合：白皮书题目本身就把问题求解作为可视分析的终极目标，并明确将“理解、推理、决策”构建为达成目标的主要认知手段。

基于复杂认知理论和可视分析的双重视角，我们可以将可视分析系统所支持的用户行为归纳为两个主要类型：问题求解和理解。这一分类既保留了复杂认知研究的理论框架，又突出了可视分析领域的研究重点。

我们在此暂时未将推理、学习和决策作为单独的分析行为来分析。这一处理方式基于以下考量：首先，在复杂认知理论中，推理被认为是一个中间过渡行为^[25]，需要依附于学习、问题求解或决策等其他认知行为而存在，起着支持其他认知行为的作用。其次，虽然可视分析领域已有支持教学行为的研究[3]，但传统上并未将学习行为本身作为核心的研究对象，这可能是因为典型的学习行为通常发生在结构化的教育场景中（如课堂、实验操作等），以知识获取为明确目标，而可视分析中涉及的“理解”行为往往是问题求解或决策的子过程。值得注意的是，近年来面向可解释人工智能的可视分析系统一些研究工作实质上就是针对的学习行为（即获取新信息、建构新知识）：例如在 Yuan 等人的一篇关于面向机器学习的可视分析研究综述中^[26]，270 篇文章中有超过一半（144 篇）涉及对模型结果的理解（understanding），然而这类研究中很少明确利用学习理论来指导用户行为分析和系统设计。

3.1.1 问题求解行为

基于问题求解的认知行为的本质是一种基于初始条件向目标结果逼近的认知过程。如果一个问

题的初始条件和目标结果都是明确而无歧义的，那么这个问题就是良构问题，否则就是非良构问题。有些良构问题是可以用算法来得到解决的。而所有的非良构问题都需要人的介入，来把模糊、有歧义的条件明确化。

可视分析系统做针对的问题通常是包含了诸多不确定性的非良构问题。以一个针对时序数据的可视分析系统为例^[27]，用户的目的是对数据作因果分析，但寻找某个现象的原因是一个非常模糊的目标，不同的假设会导致不同的分析策略。可视分析系统的目的是帮助用户分析比较不同的假设、策略。

3.1.2 决策行为

决策作为一种典型的复杂认知行为，在研究中存在两种观点：部分文献认为决策和问题求解在认知机制上有相似性^[28]，但另一些研究则强调二者本质区别^[29]。从可视分析的角度出发，我们倾向于把决策看做是一类特定的问题求解行为，因为可视分析系统中的决策行为通常表现为基于给定数据来生成候选方案，然后通过一些预设条件筛选最终方案。这个过程可以把最终方案视为一个问题的目标状态，方案生成和评估过程作为问题求解的过程。

以一个足球排阵可视分析系统为例^[30]：该系统可以帮助教练员根据某个确定的对手选定一个最佳阵容。决策过程中，系统可以根据给定信息（如对手和己方的情况）产生各种阵容，然后教练员依据一些标准（如队员的状态、之间的配合程度等）来选取最佳的阵容。这个过程有初始条件—对手和己方的各种信息、目标结果—最后的阵容，选择最佳阵容就是依据给定的条件去寻找一个解。当然这个决策问题是一个非良构问题，因为最佳阵容这个目标不是很明确，因为没有提供界定“最佳”的标准，需要教练员依靠自己的经验做出判断。

基于对 230 篇可视分析文献的分析，我们发现可视分析系统研究多聚焦于问题求解类行为

(148 篇), 因此这里我们将问题求解作为理解可视分析行为的出发点。

3.2 问题求解的相关理论

在现代科学范式下, 认知心理学已经发展出完善的问题求解理论体系^[31], 为深入研究问题求解的认知机制提供了系统的理论框架和方法论支持。

3.2.1 问题状态、问题空间、算子

一个给定的问题通常包含两个状态: 作为出发点的初始状态(通常包括给定的条件、数据)和作为目的的终结状态(即结果)。问题求解就是依据某个特定的策略, 构建一系列逻辑连贯的中间状态, 将初始状态与终结状态系统地连接起来。这些中间状态, 以及问题的初始状态和终结状态, 统称为问题状态, 而所有问题状态的集合构成了一个完整的问题状态空间。问题求解的过程可以看做是对问题空间中的中间状态进行搜索和选择的过程。通常在一个空间内存在多条合理的路径, 即多个解决问题的策略。例如从问题求解的理论看一个数学证明题时, 该题目的给定条件就是初始状态, 需要证明的结果就是终结状态, 一个合理证明的中间步骤就是中间状态。这个题可能有多种证明方法, 不同方法涉及到的中间状态可能会有所不同。该问题的初始和终结状态, 以及所有可能证明方法所涉及到的中间步骤构成了这个题目的问题空间。证明这个题目就是在该问题空间中搜索一条合乎数学逻辑和规范的路径来把相关的问题状态连接起来, 而且有效的路径往往不止一条。

在一个解决方案下, 从一个中间状态向另一个中间状态作合理地、有逻辑性地转移时, 通常需要借助于一些手段或工具, 这些手段或工具在问题求解理论中被称作算子(operators)。例如证明一个数学题需要依赖一些公理和定理, 而这些公理和定理就是解决这一问题的算子。

图 1 展示了问题状态、问题空间以及算子之间的关系。在该图中, 两条连接初始和终结状态的路

径代表了两个解决问题的策略。

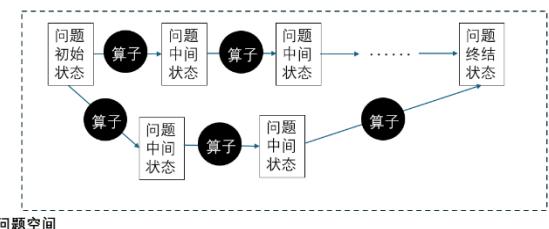


图 1 问题状态、问题空间以及算子

多数情况下, 问题求解使用已知的算子即可, 但某些情况下, 现有的算子无法满足要求, 则需要去创建新的算子。例如, 如果一个数学证明需要先证明一个新的定理, 然后在这个新定理的基础上来完成, 那么这个新定理就是一个新的算子。但新定理需要得到严格的证明, 才可以应用。

3.2.2 问题的难易、良构和非良构问题

一个问题的难易取决于多种因素。一个因素是该问题的良构性^[32]。良构问题的初始状态和终结状态明确、合理, 根据它们可以明确问题空间, 并通过确定中间状态以及相关路径来得到解决方案。课本上的数学题目往往都是良构问题, 因为其初始条件基本是充分且必要的, 目标也清晰。很多良构问题都有已知的解决方案, 可以通过工具来自动执行, 即问题求解过程的自动化。例如, 很多数学很多优化问题、几何问题都可以通过程序来解决或证明。

与之相反, 非良构问题存在着缺乏足够条件、目标模糊甚至矛盾等问题空间不确定性的特点, 造成了在确定中间状态和解决路径的困难。在这种情况下, 显然自动化的问题求解工具是无法使用的。

处理非良构问题通常需要客观主题的介入, 通过对问题的初始状态或终结状态做出一些改变来使该问题良构化, 或者在一个不确定的问题空间内做路径搜索碰到障碍时, 根据具体的情况引入局部条件和约束以完成搜索。这些处理手段往往不仅依赖于客观知识体系, 还受主观个体的经验和价值观的影响。

非良构问题因其固有的不确定性通常被视为更具有挑战性，但良构问题也可能称为困难问题。导致一个良构问题成为困难问题的原因很多，可以是一个良构问题除了初始和终结状态外，无法确定其他的中间状态，导致问题空间无法构建。当我们面对一个不知该采用如何证明一个数学题时，就陷入了这种困境。另外一种情况是即使一个问题存在着完美的空间空间并且部分中间状态也预知，但寻找一条合理的、连接这些中间状态的路径却面临困难。有时难点出现在从一个中间状态向下一个中间状态的转换上，例如在数学题目证明中时，证明的瓶颈是无法建立两个重要步骤之间的逻辑衔接。

无论是解决容易的还是困难的问题，都涉及对问题空间做搜索，而寻找探索问题空间的方法因人而异：经验丰富的专家通常采用从初始状态到终结状态的正向推理方式，而初学者则更倾向于从目标回溯的逆向推理方式^[31]。

3.2.3 问题求解的认知过程

问题求解涉及到的认知过程通常包括问题识别与表征、策略选择、执行与评估等几个关键阶段。问题求解始于对问题的识别和表征，即主观个体需要理解问题的结构和目标。根据 Newell 和 Simon 的“通用问题解决者”模型^[33]，问题表征取决于认知主体对问题初始状态、终结状态和可能路径等重要因素的了解。错误的问题表征会导致无效的解决策略，因此这一阶段至关重要。

在明确问题结构后，主观个体会选择适当的解决策略。常见的策略有系统性的、步骤清晰的解决方案和启发式的、依据经验的解决办法。例如，手段-目的分析方法 (means-ends analysis) 就是一种启发式策略，其目的是缩减当前问题状态与最终目标状态之间的差异。此外，类比推理也是一种常见的策略，这种策略借鉴过去类似问题的解决方案来指导当前的策略。

策略执行阶段涉及复杂的认知过程，主要依赖元认知 (metacognition) 来监控功能。元认知可定义为认知主体对自身认知过程的反思性认知，即

对诸如问题求解之类的复杂认知过程的觉察、监控和调控^[34]。这种元认知能力通过确保策略实施的适应性和灵活性，最终实现认知任务的有效完成。

评估与反馈阶段着眼于评估最终解决方案的有效性。若结果不符合预期，则需要重新调整问题表征或策略，形成迭代的问题求解过程。

这些问题求解理论为可视分析系统研究提供了重要的理论和方法，我们借助于问题求解的相关理论和概念（如问题状态空间、算子、非良构问题、问题求解策略等），来构建适用于可视分析的认知行为解析框架。

4 基于问题求解的可视分析研究方法

把可视分析行为看做问题求解的复杂认知行为可以帮助我们更好地理解可视分析系统研究中各个环节的目的和任务，从而完善研究方案。这里我们以可视分析系统应该有效支持问题求解为出发点，讨论相关主要研究环节需要考虑的因素。我们侧重于三个方面：用户需求分析的目的、系统设计元素的功用以及系统评估的侧重点。

4.1 需求分析：侧重问题求解的认知过程

在需求分析这个环节，无论采用什么方法，研究和设计人员必须明确需要观察和分析什么样的用户行为。目前常见的方法通常侧重于诸如用户是如何完成他们的工作的、用户在工作时现有的交互过程是怎样的、他们常遇到的问题是什么等方面。

当系统设计的目的是帮助用户解决问题，特别是非良构问题的难题时，只是关注用户的常规工作流程、交互方式等外在行为表现显然是不够的。我们需要深入了解用户在解决一个难题时与问题求解有关的认知行为和过程。虽然用户的交互行为记录（如他们用到的数据、菜单命令点击历史数据等）可以帮助我们推测其解决一个问题时所采用的基本方法，但这些交互记录信息只是用户认知过程的外在体现，但用户内部认知过程的某些重要步骤（如用户时如何确定求解策略、如何对非良构问题

进行转换、如何在不同的算子中做取舍等)却很难在这些数据中得到体现。

这就要求我们在做需求分析时要有意识地深入观察那些和问题求解有关的行为,了解他们是如何理解和诠释一个问题的(如对问题的条件和目标的分析,对问题空间的理解),如何处理非良构问题的(如对不充分条件的添加和完善、或对模糊目标的清晰化),如何筛选问题求解策略的(如对可能方案的探索、对最佳方案的筛选、对过去遇到过的类似问题的回忆)等等。

深入了解用户的内在认知过程可以采用多种研究方法。若采用问卷调查方式,则需要在问卷设计阶段对系统所针对的问题、以及解决这个问题可能碰到的诸多障碍做深入的分析,以确保问卷问题的明确性、针对性和有效性。

也可以采用更为开放的访谈调查方式,给用户自由发挥的空间来具体描述他们解决问题的过程。这种访谈方式需要访谈者首先需要对用户的工作有一定的了解,以便于能正确理解他们在问题求解过程中所采取的有效手段;其次访谈者还需要有敏锐的洞察力,能从用户的叙述中捕捉到与理解问题和解决问题相关的关键信息,并针对这些信息进行进一步跟踪、追问,以深入了解内部认知过程。

当采用现场观察法时,用户需要采用边想边说(think aloud)的方式来把指导他们交互行为的思维和认知过程做实时口头表述。通过对用户言语内容的分析,可以推断其内在的认知过程。但使用这种方法时,需要留意用户行为和语言描述之间的偏差以及操作的失误,当这类问题出现时,要及时要求用户去澄清或改正,以确保问题求解策略的正确执行。

总之,在问题求解的视角下,用户需求和分析方面的工作要着眼于那些和问题求解相关的行为和内在认知过程,力求深入了解用户如何诠释一个问题(如对问题空间的认知、对问题本质的判定)、如何制定解决方案(如对类似问题的选择、对可能备选方案的探索、对最后方案的筛选)。这些信息可以揭示用户在面对这类问题时常用的策略、方法

和工具,以及他们在相关步骤的认知资源分配。在此基础上,可以进一步总结用户的需求。

这里需要强调的是用户需求应该是在问题求解层面上的、涉及问题求解相关因素(如问题空间、策略、类比问题、算子等等)的需求。对于针对某些特定视图或交互工具的需求,不能只局限在描述用户对它们的期待,而应该进一步在问题求解的视角下进一步阐述它们在问题求解中的作用和意义。基于这种用户需求所提出的设计方案,对于提升复杂分析行为的效度会更有效,因为它们针对的是如何协助用户改善他们解决问题的方式,而不是只局限于过程中的某些具体操作。

4.2 问题求解视角下的系统设计

问题求解理论也可以帮助我们重新认识系统设计的功用,深入了解系统各部件和交互工具在问题求解过程中的作用。

4.2.1 算法/视图: 问题求解中的算子

在问题求解的视角下,我们可以把数据处理算法和视图都看做问题求解过程中必不可少的算子,并准确把握相关设计工作的作用和创新性。

算子在问题求解过程中很重要,但解决问题并非需要新的算子,除非现有的算子无法满足要求。在可视分析系统中,数据处理算法可以被看做一种算子,其功用是把数据从一个状态转换为另一个状态。对于这个环节,如果已有的算法就可以达到目的(如得到正确的转化结果),那么将该算法融入系统就足够了。

在某些情况下,当已知的算法无法满足要求时,就需要改进已有的算法、甚至创建新的算法。在问题求解的视角下,一个新算法就是一个新的算子,而新算子会对相关的问题求解过程产生影响。因此这方面工作的重要性不再局限于一个系统部件的设计创新,更在于其对一个问题求解过程的改变(如新的算子导致新的解决方案、新的算子极大地缩短了问题求解路径等),因此这类工作的评估应该从更深入的理论层次来审视其对一个问题求

解过程的影响,而不是仅仅局限于和其他类似的算法比较效率。

对于视图设计,我们也可以同样从算子的角度来理解其作用。如果已知的视图可以帮助用户从一个问题状态转移到下一个问题状态,那么无需去设计新的视图。只有当已有视图无法达到要求时,才需要去改进已有视图、或设计新的视图。同样,这种改进和创新会影响问题求解过程。

视图对问题求解的帮助时多方面的。首先,视图可以把问题求解需要的算子信息(如可用的算法和视图)或状态信息(如数据)直接展示在工作面上,这样用户就可以方便地查找相应的工具。视图在这方面功用于交互设计领域的一个共识相符:把和交互相关的一些认知信息外显化(externalization)可以减少认知过程中的记忆负担。

视图对问题求解的另外一个意义是图形可以把某些抽象的推理过程转化为空间结构的比较和变换(如使用直方图可以把抽象的数值比较简化空间物体的大小对比)。这种转化也可以影响问题求解的策略和效率。

4.2.2 交互: 支持问题求解认知过程

在解决一个问题时,我们需要相关的陈述性知识(如问题状态、算子、以往所碰到的类似问题等)和程序性知识(如对算子的使用方法、不同问题状态之间的转化方式、解决类似问题的步骤等),借助短期记忆来处理当前问题状态,依赖长期记忆来获取以往相关的类似问题和解决方案。如何建立短期记忆和长期记忆之间的联系可以通过将相关的线索可视化来达到,这就要求可视化设计要考虑到当前问题状态和可能类似问题的联系,而这种联系是基于用户的交互行为的,因此是动态的。因此系统设计应该力求帮助用户流畅地探索问题空间,从初始状态到达终结状态。

当前的交互设计多数停留在支持操作层面用户与算子的交互,而缺乏对宏观的、问题求解策略探索行为的支持。例如当面对一个非良构问题时,

用户需要了解问题空间的各个元素,以分析非良构的症结所在。而面对一个困难的良构问题时,用户可能需要了解该问题空间中的众多问题状态,以寻找下一个问题状态和可能的联通路径。在这两种情况下,用户或许期望有工具来协助他们对问题空间进行各种探索,如查看在添加了不同的初始条件后的问题空间状态的变化、检验不同的算子对问题状态之间联系的影响等等。当前的可视分析系统在这方面的工作还很少见。

问题求解过程经常会用到类比法,借鉴以往解决类似问题的策略来解决当前问题。但根据当前问题状态来回忆搜索类似问题是一个认知难题,因为很多类似的问题虽然本质一样,但他们的表面特性却不尽相同,用户需要一定的知识和经验才能透过表面现象认清问题的本质^[31]。目前的可视分析系统在如何帮助用户厘清问题的本质方面也缺乏有效的研究和工具。

另外,问题求解的策略、过程和结果与用户的知识、技能、目的等方面的相关性很强。不同的用户期望的分析工具可能差别很大,他们对问题空间的探索也不尽相同。例如,现有的认知理论表明有经验丰富的分析人员通常对相关的问题概念和状态很熟悉,所以它们侧重于依靠程序性知识来解决问题,而初学者需要先熟悉相关的概念,然后依靠概念之间的关系来寻找解决方案。因此系统在交互层面应该考虑如何支持不同用户的个性化问题求解方式,而不是给出一套标准化的工具或工作模板来让用户去适应。

阻碍问题求解的挑战有很多。对于一个非良构问题,最大的挑战是如何将其良构化。而由于每个人的知识、经验、目的等方面的不同,他们对于如何良构化同一个问题有着不同的策略,因此交互设计应该考虑帮助用户流畅地探索一个问题空间,了解问题空间的各个元素,从而帮助他们了解非良构的症结在哪里,然后决定采取什么良构化的措施(如丰富初始条件、明确最终目的等)。

对于一个良构的、但问题空间复杂或解决路径冗长的问题,交互设计除了支持用户对问题空间

的探索之外, 还应该考虑协助用户管理在路径探索过程中碰到的种种问题, 如保存重要的中间状态、展示相关状态之间的在数据和任务之间的关系等等。

4.3 系统评估: 对复杂认知的测量

基于问题求解理论, 我们也可以提升可视分析系统评估手段, 在传统的工具可用性的基础上, 考虑系统对复杂认知过程的影响。

在使用案例分析方法时, 我们可以侧重于了解用户是如何使用系统来对问题空间进行探索的。例如, 我们可以分析系统是如何帮助用户管理众多的问题状态、如何协助用户探索比较不同的问题求解路径、如何帮助用户搜索类似问题的等等。

用户访谈也应该聚焦于和问题求解的策略、方法相关的问题。例如, 访谈的问题可以涉及分析中哪一步(即一个问题状态到另一个问题状态的转化)是关键性的、借助了什么算法和视图工具、采取了什么方法给问题做了更为明确的定义(即良构化一个非良构问题)、系统所提供的工具是如何协助这些环节的等等。

除了在常见的案例分析和访谈中搜集和问题求解相关的信息外, 我们也可以通过问卷来获取相关的信息。这方面的问题应该侧重于问题求解层面的复杂认知度量, 如用户对依靠系统所得到的解决方案的信心、以及对比用户在使用系统前后对该问题的了解和熟悉程度等。

复杂认知理论也为实验验证方法提供新的度量手段。例如, Ackerman 等^[35]提出了在检验交互系统有效性的实验方法中, 可以增加诸如用户对正确与错误解决方案的辨识程度(resolution)、以及对问题解决方案的信心(confidence)等度量, 以检验系统对用户元认知能力的影响。

和传统的基于可用性度量和局限于完成某个特定任务的评估方法相比, 这些着眼于一般性的问题求解过程、元认知特征的评估手段, 可以提高研究结果的外部效度, 提升研究在理论层面的贡献。

综上所述, 以问题求解视角审视可视分析研究能够为需求分析、系统设计与评估提供更具认知深度的理论框架: 需求分析上可以突破对用户表面交互行为的观察, 深入其内在认知过程(如问题空间构建、非良构问题转化、策略筛选逻辑等); 系统设计上可以进一步明确系统中各组件对问题求解过程的支持(如数据处理算法与视图作为关键算子来启发或协助对问题状态的转换和求解路径的优化), 以及交互设计在支持动态问题空间的探索和相关策略的适配等方面的作用; 系统评估方面帮助突破传统的可用性度量, 建立面向复杂认知过程的指标体系(如系统对问题空间探索、策略筛选等关键步骤的支持能力、用户对结果的满意度和信任度等度量)。

5 其他复杂认知理论的应用

我们上一节的讨论侧重于问题求解的相关理论, 因为多数可视分析系统的研究所关注的关键问题可以归纳为问题求解类的挑战。但我们认为我们的方法可以应用于分析其他的复杂认知行为理论与可视分析行为的关系。这里我们概述决策和学习的相关理论对可视分析系统研究的指导意义。

决策行为也是认知心理学领域被深入研究的一个分支^[36]。和问题求解行为类似, 决策行为也是一个多阶段过程, 包括了问题识别与目标确定、信息搜索与选项生成、选项评估与权衡、决策执行与评估、以及结果反馈与学习等阶段, 其认知过程也是涉及元认知。通过深入理解决策理论, 我们也可以建立决策理论和分析行为的对应关系, 进而去指导面向决策行为的可视分析系统研究。我们可以根据决策的认知特性来指导需求分析, 获悉用户在决策过程所面临的挑战和哪些决策行为的要素有关(如缺乏明确的决策目的、缺乏必要的决策选项和评估标准等), 然后通过系统设计来辅助决策的关键阶段(如通过算法和视图来帮助产生所需的选项、展示关键的评估标准等)。在评估方面, 我们仍然可以采用元认知层面的度量, 来测评系统对一般性

的决策过程和元认知能力的功效。

此外, 我们还需要关注决策行为的独特特性, 探索可视分析工具对这些特性的影响。例如, 决策理论指出决策的结果会被决策者的诸多认知偏差所影响^[37], Wall^[38]等也分析了在可视分析过程中的一些主要认知偏差。基于这种理解, 系统的设计在支持决策要素和决策过程的同时, 也需要考虑如何通过技术手段来降低或减免认知偏差对决策的影响(如设计算法来捕捉分析过程中的认知偏差, 用视图展现认知偏差可能产生的决策错误等), 从而让用户能察觉到潜在的认知偏差, 及时纠正, 做到正确决策。

同样, 我们也可以依据和学习相关的理论(如自我调节学习理论课, self-regulated learning)^[39]来指导和学习、理解相关的可视分析系统研究。自我调节学习所涉及的认知活动也是一个多阶段的过程, 也包含了元认知行为。面向学习和理解行为的可视分析系统可以根据学习过程以及相关元认知的特点, 来分析用户的行为, 设计适当的工具来提升学习的功效。例如, 当前针对可解释人工智能的可视分析系统研究大多侧重于对算法参数、模型过程等客观因素的呈现(这方面工作的综述见^[26]), 而忽略了用户本身和学习行为相关的主观因素(如先验知识水平、知识构建方式等)在学习过程中的作用, 以及主客观因素之间的动态交互机制。这些方法上的局限性制约了此类系统在促进学习行为方面的实际效能。在学习理论的框架下, 这类系统的研究可以着眼于了解用户在学习人工智能模型过程中的目标、方法和障碍, 设计基于学习理论的视图和交互手段来帮助用户有效地掌握模型中的关键概念、指标、工作流程等, 然后通过测量用户在使用系统前后的知识结构的变化来验证系统的有效性^[40]。

6 结语

我们通过分析当前可视分析系统研究中的主流方法, 指出这些基于简单认知行为的方法与面向

复杂分析行为的研究目标存在着错配, 并以复杂认知领域的问题求解理论为例, 阐述了如何建立复杂认知理论与可视分析行为的关系, 以及这些理论对可视分析研究主要手段的指导意义。

总之, 我们认为复杂认知理论与可视分析实践的深度融合可以帮助我们探索可视分析系统研究的新范式。通过深入了解问题求解、决策和学习等领域中和复杂认知相关的理论, 我们可以构建面向可视分析系统设计的、具有更高内部效度和外部效度的方法论体系。

参考文献(References):

- [1] D. Deng, J. Wu, and J. Wang, ‘Eventanchor: Reducing human interactions in event annotation of racket sports videos’, *Conference on Human Factors in Computing Systems - Proceedings*, 2021, doi: 10.1145/3411764.3445431.
- [2] Y. Hou *et al.*, ‘Visual Evaluation for Autonomous Driving’, *IEEE Trans Vis Comput Graph*, vol. 28, no. 1, pp. 1030–1039, 2022, doi: 10.1109/TVCG.2021.3114777.
- [3] J. S. Wong and X. “Luke” Zhang, ‘Message-Lens: A Visual Analytics System to Support Multifaceted Exploration of MOOC Forum Discussions’, *Visual Informatics*, vol. 2, no. 1, pp. 37–49, 2018, doi: 10.1016/j.visinf.2018.04.005.
- [4] W. Yang *et al.*, ‘Diagnosing Concept Drift with Visual Analytics’, *Proceedings - 2020 IEEE Conference on Visual Analytics Science and Technology, VAST 2020*, pp. 12–23, 2020, doi: 10.1109/VAST50239.2020.00007.
- [5] L. Gou, X. Zhang, A. Luo, and P. F. Anderson, ‘SocialNetSense: Supporting sensemaking of social and structural features in networks with interactive visualization’, *IEEE Conference on Visual Analytics Science and Technology 2012, VAST 2012 - Proceedings*, pp. 133–142, 2012, doi: 10.1109/VAST.2012.6400558.
- [6] L. Gou and X. L. Zhang, ‘Treenetviz: Revealing patterns of networks over tree structures’, *IEEE Trans Vis Comput Graph*, vol. 17, no. 12, pp. 2449–2458, 2011, doi: 10.1109/TVCG.2011.247.
- [7] Y. Wang, Y. Qian, X. Qi, N. Cao, and D. Wang, ‘InnovationInsights: A Visual Analytics Approach for Understanding the Dual Frontiers of Science and Technology’, *IEEE Trans Vis*

- [8] *Comput Graph*, vol. 30, no. 1, pp. 518–528, 2024, doi: 10.1109/TVCG.2023.3327387.
- [9] B. Schneiderman, ‘The Eyes Have It : A Task by Data Type Taxonomy for Information Visualizations Ben Shneiderman University of Maryland’, *IEEE Symposium on Visual Languages*, 1996.
- [10] P. Pirolli and S. Card, ‘The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis’, *Proceedings of International Conference on Intelligence Analysis*, vol. 2005, no. May 2014, pp. 2–4, 2005, [Online]. Available: <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+Sensemaking+Process+and+Leverage+Points+for+Analyst+Technology+as+Identified+Through+Cognitive+Task+Analysis#0>
- [11] R. K. Yin, *Case study research : design and methods / Robert K. Yin*. 2014.
- [12] A. Newell, ‘Unified Theories of Cognition (William James Lectures)’, p. 576, 1994, [Online]. Available: <http://www.amazon.com/Unified-Theories-Cognition-William-Lectures/dp/0674921011>
- [13] G. G. Robertson, S. K. Card, and J. D. Mackinlay, ‘Information visualization using 3D interactive animation’, *Commun ACM*, vol. 36, no. 4, 1993, doi: 10.1145/255950.153577.
- [14] A. Newell and S. K. Card, ‘The Prospects for Psychological Science in Human-Computer Interaction’, *Hum Comput Interact*, vol. 1, no. 3, pp. 209–242, 1985, doi: 10.1207/s15327051hci0103_1.
- [15] K. Frankish, ‘Dual systems and dual attitudes’, *Mind & Society*, vol. 11, no. 1, pp. 41–51, Jun. 2012, doi: 10.1007/s11299-011-0094-5.
- [16] J. R. Lewis, ‘The System Usability Scale: Past, Present, and Future’, *Int J Hum Comput Interact*, vol. 34, no. 7, 2018, doi: 10.1080/10447318.2018.1455307.
- [17] D. R. Olsen, ‘Evaluating user interface systems research’, *UIST: Proceedings of the Annual ACM Symposium on User Interface Software and Technology*, pp. 251–258, 2007, doi: 10.1145/1294211.1294256.
- [18] C. Quintão, P. Andrade, and F. Almeida, ‘How to Improve the Validity and Reliability of a Case Study Approach?’, *Journal of Interdisciplinary Studies in Education*, vol. 9, no. 2, pp. 273–284, 2020, doi: 10.32674/jise.v9i2.2026.
- [19] A. M. Riege, ‘Validity and reliability tests in case study research: A literature review with “hands-on” applications for each research phase’, *Qualitative Market Research: An International Journal*, vol. 6, no. 2, pp. 75–86, 2003, doi: 10.1108/13522750310470055.
- [20] J. Preece, *Interaction Design - Beyond Human-Computer Interaction (1st edition)*. 2002.
- [21] J. J. Thomas and K. A. Cook, ‘Illuminating the path: The research and development agenda for visual analytics’, *IEEE Computer Society*, 2005.
- [22] K. J. Keim Daniel and G. rey E. and F. Mansmann, *Mastering the Information Age Solving Problems with Visual Analytics*. 2010. doi: 10.1016/j.procs.2011.12.035.
- [23] M. Knauff and A. G. Wolf, ‘Editorial: Complex cognition: The science of human reasoning, problem-solving, and decision-making’, *Cogn Process*, vol. 11, no. 2, pp. 99–102, 2010, doi: 10.1007/s10339-010-0362-z.
- [24] J. Funke, ‘Complex problem solving: A case for complex cognition?’, *Cogn Process*, vol. 11, no. 2, pp. 133–142, 2010, doi: 10.1007/s10339-009-0345-0.
- [25] D. Dörner and J. Funke, ‘Complex Problem Solving : What It Is and What It Is Not’, *Front Psychol*, vol. 8, no. July, pp. 1–11, 2017, doi: 10.3389/fpsyg.2017.01153.
- [26] J. P. Leighton, ‘Defining and Describing Reason’, *The Nature of Reasoning*, pp. 3–11, 2011, doi: 10.1017/cbo9780511818714.001.
- [27] J. Yuan, C. Chen, W. Yang, M. Liu, J. Xia, and S. Liu, ‘A survey of visual analytics techniques for machine learning’, 2021. doi: 10.1007/s41095-020-0191-7.
- [28] Z. Deng *et al.*, ‘Compass: Towards Better Causal Analysis of Urban Time Series’, *IEEE Trans Vis Comput Graph*, vol. 28, no. 1, pp. 1051–1061, 2022, doi: 10.1109/TVCG.2021.3114875.
- [29] J. Adair, *Decision Making and Problem Solving Strategies (Sunday Times Creating Success)*. 2012.
- [30] H. A. Simon *et al.*, ‘INFORMS Interfaces’, *Interfaces (Providence)*, vol. 17, no. 5, 1987.
- [31] A. Cao *et al.*, ‘Team-Builder: Toward More Effective Lineup Selection in Soccer’, *IEEE Trans Vis Comput Graph*, vol. 29, no. 12, 2023, doi: 10.1109/TVCG.2022.3207147.
- [32] M. Bassok and L. R. Novick, ‘Problem Solving’, in *The Oxford Handbook of Thinking and Reasoning*, Oxford University Press, 2012. doi: 10.1093/oxfordhb/9780199734689.013.0021.
- [33] S. K. Reed, ‘The Structure of Ill-Structured (and Well-Structured) Problems Revisited’, 2016. doi: 10.1007/s10648-015-9343-1.
- [34] J. C. S. Newell A. and H. A. Simon, ‘Report on a general problem-solving program’, in *Proc. Internat'l Conf. on Information*

- Processing*, 1959.
- [34] C. Hertzog and R. Dixon, *Metacognition : knowing about knowing / edited by Janet Metcalfe and Arthur P. Shimamura*. 1994. [Online]. Available: <http://hollis.harvard.edu/?q=isbn:0-262-13298-2>
- [35] R. Ackerman, A. Parush, F. Nassar, and A. Shtub, ‘Metacognition and system usability: Incorporating metacognitive research paradigm into usability testing’, *Comput Human Behav*, vol. 54, pp. 101–113, 2016, doi: 10.1016/j.chb.2015.07.041.
- [36] M. Knauff and A. G. Wolf, ‘Editorial: Complex cognition: The science of human reasoning, problem-solving, and decision-making’, *Cogn Process*, vol. 11, no. 2, pp. 99–102, 2010, doi: 10.1007/s10339-010-0362-z.
- [37] D. Kahneman, *Thinking , Fast and Slow* (Abstract). 2011.
- [38] E. Wall, L. Blaha, C. Paul, K. Cook, A. E.-C. biases in, and undefined 2018, ‘Four perspectives on human bias in visual analytics’, *Springer*, Accessed: Oct. 31, 2018. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-95831-6_3
- [39] R. A. Bjork, J. Dunlosky, and N. Kornell, ‘Self-regulated learning: Beliefs, techniques, and illusions’, *Annu Rev Psychol*, vol. 64, pp. 417–444, 2013, doi: 10.1146/annurev-psych-113011-143823.
- [40] S. Ohlsson, ‘Complex Declarative Learning’, in *Encyclopedia of the Sciences of Learning*, 2012. doi: 10.1007/978-1-4419-1428-6_295.