

跨学科知识组织的概念关联研究*

王忠义 彭思源 夏立新

摘 要 大科学时代,跨学科研究成为科学发展的重要方向,这给跨学科知识服务带来了重要机遇和严峻挑战。为提供跨学科知识服务,图书情报人员需要对不同学科知识进行跨学科组织,而学科概念跨学科关联是实现跨学科知识组织的关键。为此,本文首先从概念归纳和概念移植两种概念生成方式出发,对学科概念跨学科关联的形成机制进行分析,为跨学科概念关联发现提供理论支撑;而后,从语法—语义—语用三个层面构建跨学科概念关联发现的概念表示模型,充分揭示学科概念的特征,实现学科概念的全息表达;接着,基于学科概念表示模型,从语法关联、语义关联和语用关联三个维度探讨学科概念跨学科关联的实现方法,并在此基础上归纳出学科概念跨学科关联的七种方式,为实现跨学科知识组织奠定基础;最后,通过案例分析验证了学科概念表示模型、学科概念跨学科关联方法和学科概念跨学科关联方式在学科概念跨学科关联发现中具有跨学科性、多维性、学科针对性、语义性和系统性等优势,这对于支撑跨学科知识组织具有一定程度的参考价值。图9。表2。参考文献40。

关键词 知识组织 跨学科 概念关联 概念表示

分类号 G254.0

The Concept Correlation of Interdisciplinary Knowledge Organization

WANG Zhongyi, PENG Siyuan & XIA Lixin

ABSTRACT

In big science era, with the major issues in the fields of economy, science and technology, and ecology becoming more and more complex and comprehensive, the limitations of single discipline knowledge to solve such problems are gradually exposed. In this context, there is an increasingly urgent need to break through discipline barriers and integrate different discipline knowledge to carry out interdisciplinary research. Therefore, interdisciplinary research has gradually become an important trend of scientific research, which brings important opportunities and severe challenges to interdisciplinary knowledge services. To provide interdisciplinary knowledge services, library and information staff urgently need to organize knowledge of different disciplines. As the key to achieving interdisciplinary knowledge organization, interdisciplinary

* 本文系国家社会科学基金重大项目“新时代我国文献信息资源保障体系重构研究”(编号:19ZDA345)和教育部人文社会科学研究规划基金“跨学科知识组织中学科概念跨学科关联研究”(编号:21YJA870003)的研究成果之一。(This article is an outcome of the major project “Research on the Reconstruction of China’s Document Information Resources Guarantee System in the New Era”(No. 19ZDA345) supported by the National Social Science Foundation of China and the planning fund project “Research on Interdisciplinary Correlation of Disciplinary Concepts in Interdisciplinary Knowledge Organization”(No. 21YJA870003) supported by the Humanities and Social Science Research Foundation of Ministry of Education of China.)

通信作者:王忠义,Email:wzywzy13579@163.com,ORCID:0000-0001-8945-783X (Correspondence should be addressed to WANG Zhongyi,Email:wzywzy13579@163.com,ORCID:0000-0001-8945-783X)

concept correlation ought to be attached great importance. At present, although some concept correlation methods have been proposed, such as co-word analysis, knowledge mining method, and literature-related discovery, there are still some difficulties in using these methods to realize interdisciplinary concept correlation.

To this end, this paper first analyzes the formation mechanism of interdisciplinary concept correlation from the perspective of the ways of discipline concept generation. By analyzing the formation mechanism of interdisciplinary concept correlation in the ways of concept induction, interdisciplinary concept deduction, and interdisciplinary concept metaphor, it is concluded that interdisciplinary concept correlation is mainly reflected in grammar, semantics, and pragmatics, which provides theoretical support for interdisciplinary concept correlation discovery. Secondly, based on the conclusion above, this paper constructs the concept representation model of interdisciplinary concept correlation discovery from the three aspects of grammar, semantics, and pragmatics, which not only fully reveals the characteristics of subject concepts but also realizes the deep semantic mining of concepts. Thirdly, this paper discusses the methods of interdisciplinary concept correlation discovery from the three dimensions of grammatical similarity calculation, semantic similarity calculation, and pragmatic similarity calculation. Based on the results of interdisciplinary concept correlation, the paper summarizes seven ways of interdisciplinary concept correlation according to whether the concepts are related in the three levels of grammar, semantics, and pragmatics, which contributes to laying a foundation for the realization of interdisciplinary knowledge organization. Finally, from China's discipline system, discipline pairs and interdisciplinary concept pairs under these discipline pairs are randomly selected to verify the validity and usability of the discipline concept representation model and the interdisciplinary concept correlation methods proposed in this paper. As a result, the case analysis shows that the discipline concept representation model, interdisciplinary concept correlation methods, and ways of interdisciplinary concept correlation proposed in this paper have the advantages of interdisciplinary, multi-dimensional, disciplinary specific, semantic, and systematic in the interdisciplinary concept correlation, which are of great significance and value for supporting interdisciplinary knowledge organization. 9 figs. 2 tabs. 40 refs.

KEY WORDS

Knowledge organization. Interdisciplinary. Concept correlation. Concept representation.

0 引言

在当今大科学时代,许多科学问题(如经济、科技和生态等领域的复杂问题)的范围、规模不断扩大,呈现出高度综合性特征。在这一背景下,单一学科知识的局限性逐渐暴露出来,跨学科研究成为当代科学发展的重要议题。随着跨学科研究逐渐增多,科研人员对跨学科知识服务的需求也越来越强烈,这给图书情报学

带来严峻挑战。图书情报人员需要对不同学科知识进行跨学科组织,以突破学科壁垒,整合不同学科知识,满足人们的跨学科知识需求。知识组织的基础是知识组织系统(如叙词表、本体、语义网等),若要实现跨学科知识组织,就需首先构建跨学科的知识组织系统。知识组织系统的核心在于概念及概念间关联的描述和揭示^[1],有效挖掘学科概念间的跨学科关联就成为实现跨学科知识组织的关键。当前已经存在一些概念关联的方法,如文献计量的方法、知识

挖掘的方法、非相关文献知识发现方法等,然而在借助这些方法实现跨学科概念关联时仍存在一些困难,究其原因在于,不同学科之间在研究范式、学科文化等方面存在严重隔阂,各学科都构建了自己的术语体系,由此形成的一个个学科知识孤岛,给学科概念的跨学科关联带来了严重阻碍。可见,若要实现跨学科知识组织,急需打破学科隔阂,深入剖析学科概念特征,探索学科概念的跨学科关联问题。为此,本文对学科概念的跨学科关联问题进行研究,以期对跨学科知识组织提供支撑。

1 研究综述

揭示概念关联是知识组织研究的起点^[2]。依据概念关联方法的适用范围不同,已有的概念关联方法可划分为通用概念关联方法和跨学科概念关联方法。

1.1 通用概念关联方法

依据方法性质的不同,通用概念关联方法又可细分为基于文献计量的概念关联和基于相似匹配的概念关联。

1.1.1 基于文献计量的概念关联

基于文献计量的概念关联主要通过共词分析来实现,基本思想是通过统计文献集中两两概念的共现情况来反映其关联强度。如 Rokaya 等^[3]采用共词分析揭示领域内术语之间的联系;王立学等^[4]以情报学领域为例,通过关键词共现来揭示领域内研究主题(关键词)之间的关联关系;叶鹰等^[5]利用关键词共现揭示概念之间的关联关系。基于文献计量的概念关联方法虽然可操作性强,但是该方法较少从语义关联的角度揭示概念间的关联,准确性低,所以采用该方法进行学科概念跨学科关联分析时存在一定的局限性。

1.1.2 基于相似匹配的概念关联

基于相似匹配的概念关联主要借助知识挖掘方法,通过概念相似度计算来实现,经历了基

于字面匹配的概念关联和基于概念匹配的概念关联两个阶段。

(1) 基于字面匹配的概念关联

基于字面匹配的概念关联通过计算概念间名称的相似度来揭示概念是否相关,比较经典的算法有编辑距离法^[6]、Jaccard 系数^[7]等。其中,编辑距离法通过衡量一个字符串转换为另一个字符串的最少操作次数来计算概念名称相似度;Jaccard 系数通过计算两个概念字符串集合的交集和并集的比值大小来得到概念相似度。然而,由于名称相同的学科概念在不同学科中的含义可能不同,或者不同学科中语义相似的学科概念名称也不一定相似,因此,用基于字面匹配的方式来发现学科概念的跨学科关联缺乏一定的准确性。

(2) 基于概念匹配的概念关联

基于概念匹配的概念关联强调对概念语义的理解,可细分为基于语料库的概念关联、基于语义词典的概念关联和基于混合资源的概念关联三类^[8-11]。

基于语料库的概念关联通过概念上下文信息的概率分布来计算概念语义相似度。如 Salton^[12]提出了词袋法(Bag of Words),该方法将词汇上下文中其他词汇的概率分布表示为词汇情境向量,通过计算向量间的夹角余弦值得到词汇语义相似度;Deerwester 等^[13]提出的潜在语义分析法(Latent Semantic Analysis)是对词袋法的改进,该方法通过构建词汇—文档矩阵来降低原始向量空间的噪音;韩普等^[14]基于中文大规模语料库构建了词汇共现网络,并在复杂网络的基础上进行词汇相似度计算。基于语料库的概念关联方法能够通过概念上下文情境反映概念语义相似度,但是仅通过上下文信息很难反映学科概念的学科特性,因此在学科概念的跨学科关联发现中具有局限性。

基于语义词典的概念关联主要依据语义词典所描述的概念上下位关系、同义关系等计算概念语义相似度,主要包括基于语义距离的方法、基于特征匹配的方法、基于信息量的方法和

混合方法四种类型。具体而言,基于语义距离的方法依据概念在语义词典中的路径长度来计算相似度,如 Rada 方法^[15];基于特征匹配的方法依据概念的特征如性质、作用等属性的重合度来计算概念相似度,较为经典的是 Tversky 方法^[16];基于信息量的方法依据两个概念共享信息的多少来计算相似度,如 Resink 方法^[17];混合方法是指将上述单一方法混合起来,具有代表性的是 Rodriguez 提出的 MD 模型^[18]。基于语义词典的概念关联方法能准确反映出概念语义的相关性,但揭示的概念关联是普遍意义上的语义关联,不具有学科针对性,所以利用该方法进行跨学科概念关联分析存在一定的局限性。

由于单一的概念关联方法都存在一定的弊端,所以基于混合资源的概念关联方法应运而生,如 Li 等^[19]探讨了基于多种信息源计算词汇语义相似性的算法,张硕望等^[20]将知网与搜索引擎融合起来计算词汇语义相似度,吴克介和王家伟^[21]提出一种融合知网和搜索引擎的词汇语义相似度计算方法。基于混合资源的概念关联方法在概念相似度计算上能够得到更加客观准确的结果,但是它所发现的语义关联仍然不具有学科针对性。

1.2 跨学科概念关联方法

跨学科的概念关联方法是指揭示不同学科概念间关联关系的专门方法,具体表现为非相关文献知识发现方法。非相关文献知识发现方法是 Swanson 教授^[22]于 1986 年创建的,旨在从表面上不存在任何联系的独立学科文献间发现跨学科的隐含未知的联系^[23-24]。依据知识发现的模式,非相关文献知识发现方法可分为基于非相关文献二阶知识发现和基于非相关文献三阶知识发现的跨学科概念关联方法。

1.2.1 基于非相关文献二阶知识发现的跨学科概念关联

基于非相关文献二阶知识发现的跨学科概

念关联建立在 Swanson 提出的 ABC 模式上,基本思想为:用 A、B、C 代表三个概念,若已经在发表的某学科文献中表明 A 与 B 存在关联,而在另一个学科已发表的文献中表明 B 和 C 存在关联,由逻辑递推关系可知 A 和 C 之间可能存在隐含联系,且这种关联事先并未被任何文献所记载。Swanson 最早将这一模式应用至生物学领域,发现了“雷诺氏病”与“鱼油”、“偏头痛”与“镁”两对跨学科概念之间的关联性^[25,26];我国学者黄水清^[27,28]等为验证非相关文献知识发现方法在中文文献中的适用性,先后采用闭合式和开放式非相关文献知识发现方法在中文文献中重现“雷诺氏病”与“鱼油”、“偏头痛”与“镁”两对概念的知识发现过程;刘小慧等^[29]运用闭合式非相关文献知识发现方法挖掘潜在的跨学科合作研究主题。基于非相关文献的二阶知识发现是目前跨学科概念关联的主流模式^[30],但是该方法不能完全反映出概念的语义内涵,准确性有待提高。

1.2.2 基于非相关文献三阶知识发现的跨学科概念关联

非相关文献的三阶知识发现方法由我国学者李勇等^[31]提出,基本思想为:用 A、B、C、D 代表四个概念,若 A 和 B、B 和 C、C 和 D 分别在不同组的文献中共现,由逻辑递推关系可知 A 和 D 之间可能存在隐含联系,且尚未有文献记载 A 和 C、A 和 D、B 和 D 之间的关系。基于非相关文献三阶知识发现的跨学科概念关联能够挖掘出许多基于非相关文献二阶知识发现无法发现的潜在关联,但其仍无法揭示概念语义内涵,因此存在局限性。

综上,目前通用的概念关联方法和已有的跨学科概念关联方法在实现学科概念跨学科关联时均存在局限性,具体表现为:①通用的概念关联方法缺乏学科针对性,它主要通过概念的名称相似或语义关联来揭示相关概念,揭示的关联只是从概念一般特征的角度出发,缺乏对概念学科属性、学科使用情境等学科特征的揭示,而跨学科的概念关联发现须以学科为参照

系,充分结合概念的学科特征,发现更为具体的概念关联;②已有的跨学科概念关联方法缺乏语义深度挖掘,它主要是从文献的角度对不同学科潜在的相关概念进行揭示,没有完全反映概念自身的语义特征,因此无法实现对不同学科概念关联的准确揭示。针对以上问题,本文对学科概念的跨学科关联进行研究,以支撑跨学科知识组织的深入发展。

2 学科概念的跨学科关联

学科概念的跨学科关联旨在揭示不同学科

的相关概念,建立学科概念间的跨学科联系。为此,本文提出学科概念跨学科关联的研究框架(见图1)。具体而言,首先从概念归纳和概念移植两种概念生成方式出发,分析学科概念跨学科关联的形成机制;接着,在学科概念跨学科关联形成机制的指导下,基于对学科概念的语法分析、语义分析和语用分析构建学科概念表示模型;然后在学科概念表示模型的基础上从学科概念的语法、语义和语用三个层次探讨学科概念跨学科关联的实现方法,并在此基础上总结学科概念跨学科关联方式的复杂类型,以期为跨学科知识组织提供支撑。

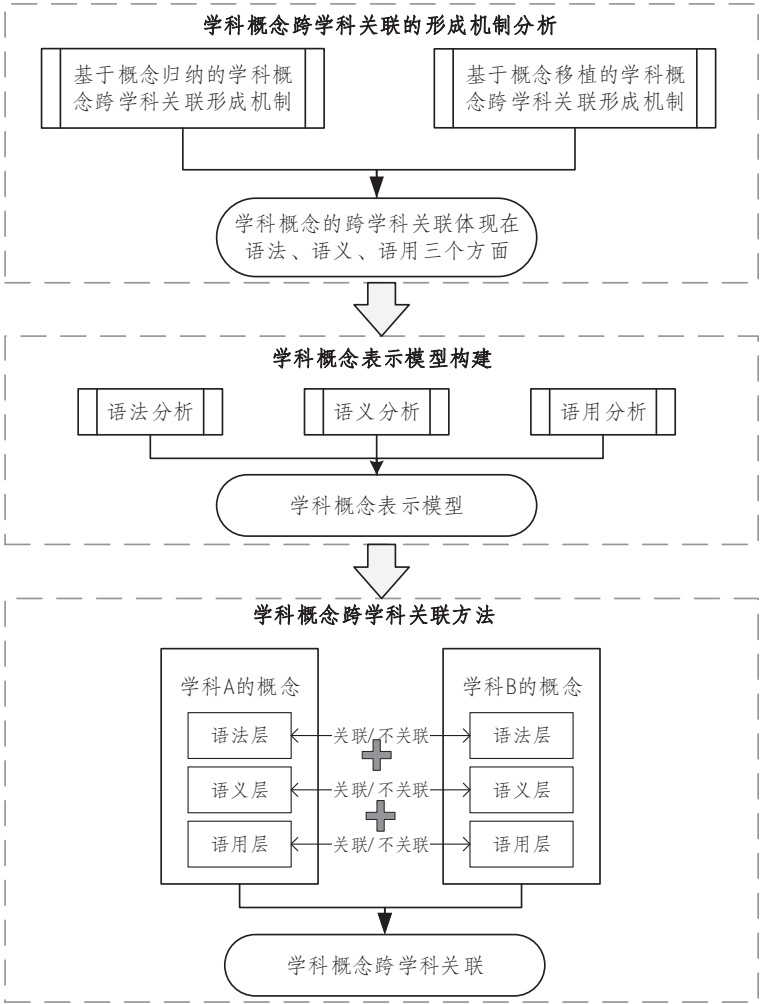


图1 学科概念跨学科关联研究框架

2.1 学科概念跨学科关联的形成机制分析

在对学科概念进行跨学科关联时,若要真正打破学科藩篱,建立学科概念间的跨学科联系,就需要探讨学科概念跨学科关联的形成机制。而学科概念跨学科关联的形成机制又与学科概念生成方式密不可分,学科概念的生成方式包括归纳和移植两种,在不同的概念生成方式下学科概念跨学科关联的形成机制不同。为此,本文分别基于归纳和移植两种概念生成方式对学科概念跨学科关联的形成机制进行分析,以期对学科概念跨学科关联提供理论支撑。

2.1.1 基于概念归纳的学科概念跨学科关联形成机制

归纳作为学科概念生成的一种重要方式,是指某学科通过对学科研究对象的观察形成经验性认识,在此基础上,抽象概括出实例对象的特征属性,最后以学科概念的形式把这种思维结晶固定下来的过程。由此可见,基于归纳的学科概

念生成过程具有相对的学科独立性,那么在这一情境下学科概念中的跨学科关联是如何形成的?通过深入分析发现,之所以不同学科通过各自归纳形成的学科概念间仍存在跨学科关联,主要原因在于学科概念所反映的实例对象间存在关联性,也即基于概念归纳的跨学科概念关联主要是由于不同学科实例对象间存在相似性和相关性(见图2)。由于不同学科研究的实例对象具有关联性,因此尽管从不同学科实例对象抽象出的概念的术语名称不同(也即术语异构),但是概念内涵(也即属性)却存在跨学科关联。可见,基于归纳的学科概念跨学科关联主要体现在概念属性和概念实例对象的跨学科关联上。例如植物学和动物学研究的实例对象植物和动物在生长规律上具有相似性,因此,通过各自归纳形成的“发芽”概念和“出生”概念,虽然术语名称不同,但都表征了一种新生的状态属性,体现出学科概念跨学科关联的特征。

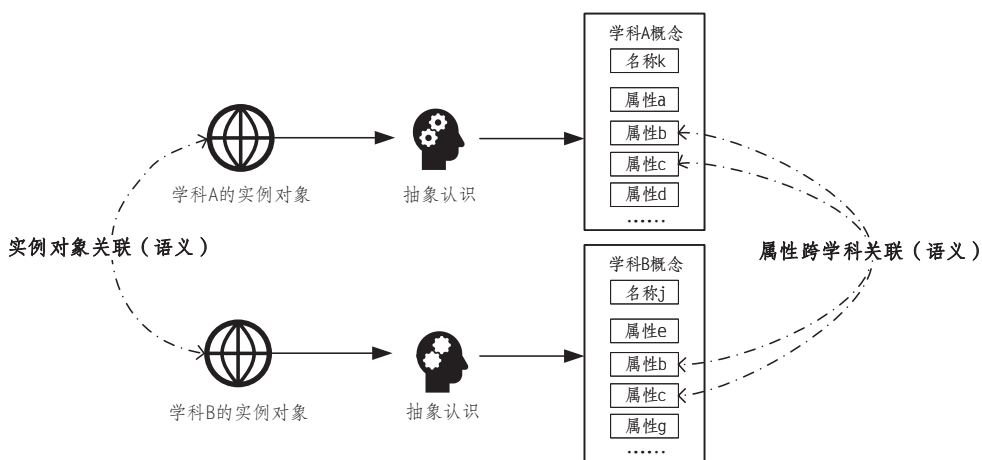


图2 基于概念归纳的学科概念跨学科关联形成机制

2.1.2 基于概念移植的学科概念跨学科关联形成机制

在科学研究中,我们把学科概念间的交流,即由一个学科引入至另一个学科的现象叫作跨学科概念移植。跨学科概念移植在概念生成中被广泛应用,是学科知识创新的重要方法,其优势在于用熟悉的概念说明新现象和新知识,从

而为人们提供一个易于理解的概念框架,简化概念生成的复杂程度。跨学科概念移植能够直接反映学科概念间的跨学科关联,具体包括跨学科概念演绎和跨学科概念隐喻两种不同的移植方式,在不同的移植方式下,学科概念跨学科关联的形成机制不同,学科概念跨学科关联的表现也不同。

(1) 基于跨学科概念演绎的学科概念跨学科关联形成机制

跨学科概念演绎是指基于学科概念的本质内涵,学科概念演绎至其他学科产生新知识的一种方式。在学科概念的跨学科演绎过程中,新学科在继承原概念的本质属性的同时,会根据学科自身需求对原概念进行适应性改造,赋予原概念新的学科内涵以形成新概念,且原概念和新概念在概念术语名称上也呈现出演化现象,能够反映学科概念跨学科演绎的特征(见图3)。由此可见,基于跨学科概念演绎的学科概念跨学科关联主要是由于原概念与新概念之间

的本质的同一性导致的,跨学科概念关联相应地体现为概念名称的跨学科关联和概念属性的跨学科关联,且学科概念的跨学科关联不仅存在于原概念与新概念之间,也存在于同源演绎出的新概念之间。在此本文以“熵”的跨学科演绎过程为例,“熵”最初起源于热力学,表达运动的无序性和混乱性,后基于这一本质意义在多个学科进行了跨学科概念演绎,如信息学中的“信息熵”、生物学中的“负熵”和数学中的“测度熵”等概念,它们在概念名称和属性上都具有一定的跨学科关联性。

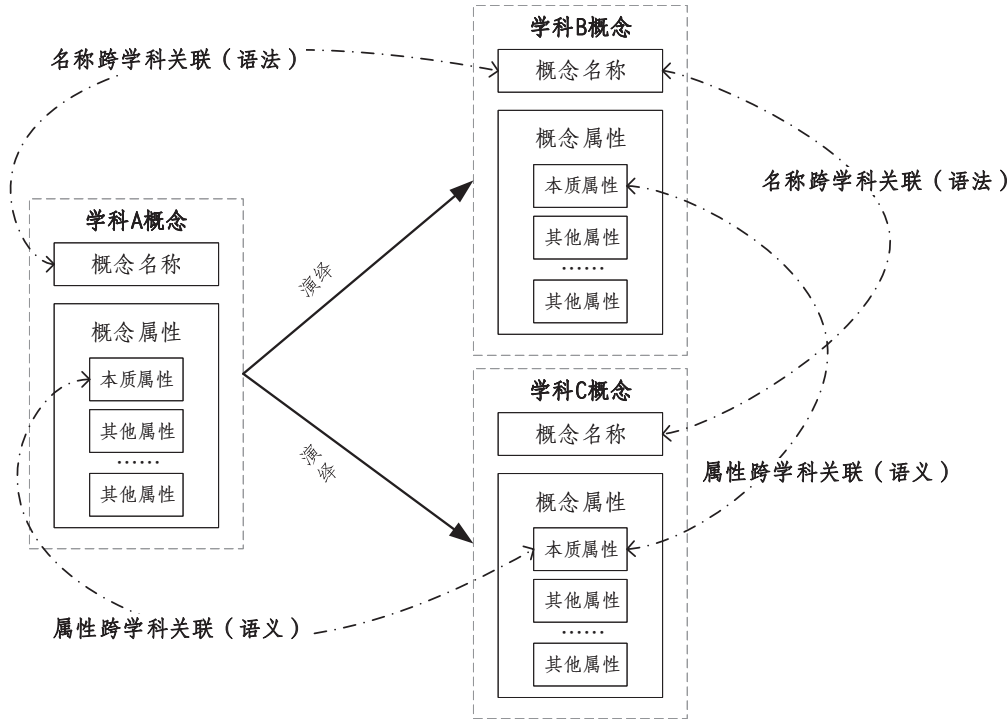


图3 基于跨学科概念演绎的学科概念跨学科关联形成机制

(2) 基于跨学科概念隐喻的学科概念跨学科关联形成机制

跨学科概念隐喻是指通过对一个学科(源域)概念的理解去认知另一个学科(目标域)的事物,进而建构出新概念的一种思维方式和认知手段。它具有两个显著特征:首先,跨学科概念隐

喻的发生离不开特定的情境空间,只有在特定情境之中,才能顺利将源域的概念一一映射至目标域,从而建构新学科的概念来认识新事物;其次,跨学科概念隐喻以不同学科概念之间的相似特征(也即属性)作为依据进行“跨域映射”,实现不同学科概念的跨学科关联。参与映射的往往只

是源域概念的部分特征,而源域特征的选择则依赖于特定的情境空间,在特定情境之中,源域的特征一一映射至目标域,从而建构新学科的概念来认识新事物。由此可见,基于跨学科概念隐喻的学科概念跨学科关联是在特定的情境空间中生成的一种跨学科语义映射,因而,基于跨学科概念隐喻的学科概念跨学科关联主要表现为概念内涵的跨学科关联和概念情境的跨学科关联(见图4)。具体而言,跨学科概念隐喻通过概念属性的映射以使目标域的概念更加明晰,进而实现概念属性的跨学科关联。在跨学科概念隐喻

中,情境空间主要表现为研究对象的研究主题情境,实际上研究主题是没有学科界限的,多个学科可能从事同一个主题的研究,这为学科概念的跨学科关联提供了可能性。例如计算机科学在认识“计算机”这一概念时,将其隐喻为生命科学的“大脑”,从而在“信息处理”这一主题域下,实现了跨学科隐喻概念对“大脑—计算机”之间“记忆—存储”“记住—写入”“回忆—可读”等属性的跨域映射,从而建立起生命科学的“大脑”与计算机科学的“计算机”两个学科概念之间的跨学科关联。

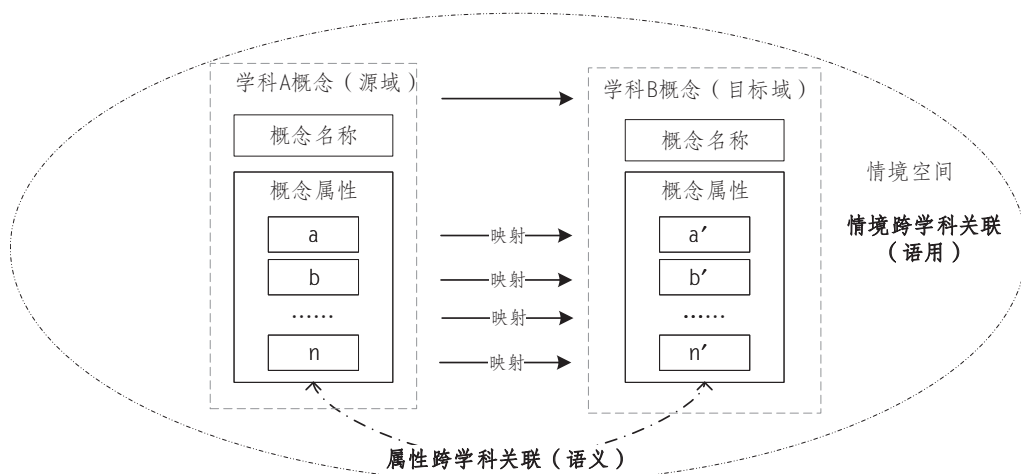


图4 基于跨学科概念隐喻的学科概念跨学科关联形成机制

综上所述,基于概念归纳的学科概念跨学科关联表现为概念属性和概念外延的跨学科关联,基于跨学科概念移植的学科概念跨学科关联表现为概念名称、属性和情境的跨学科关联。由此可见,学科概念跨学科关联主要表现在学科概念名称、属性、外延和情境上,也即概念的语法、语义(包括属性、外延)和语用三个层面。基于上述分析,为了全面且有效地发现学科概念的跨学科关联关系,下文从语法、语义和语用三个层面对学科概念跨学科关联中的学科概念表示模型进行构建,为跨学科概念关联方法的提出奠定基础。

2.2 学科概念表示模型构建

学科概念表示是使概念结构化、符号化的过程,在学科概念跨学科关联中起着重要作用。由前文可知,学科概念的跨学科关联主要体现在语法、语义和语用上,为此,本文从语法—语义—语用三个层面进行学科概念表示(见图5),以期实现对学科概念的全息表达。具体而言,通过对学科概念的语法分析、语义分析和语用分析,得到学科概念在语法层、语义层和语用层的具体描述内容,而后集成为学科概念描述框架。其中,语法层与学科概念的形式结构直接相关;语义层研究学科概念的内涵意义,在全面揭示概念学科特征的同时实现概念语义的深度

挖掘;语用层则侧重于研究学科概念的使用情境,也是概念学科特征的体现。

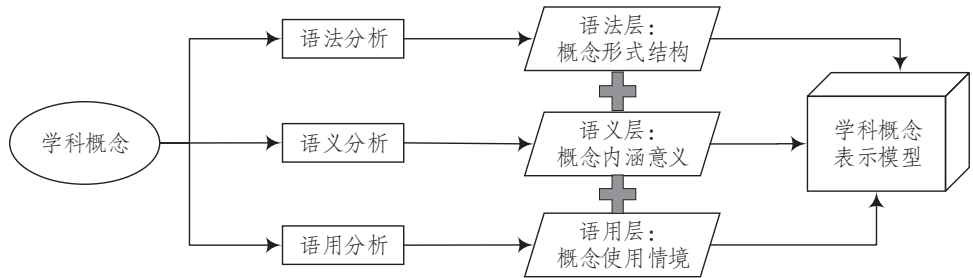


图5 学科概念表示的分析框架

2.2.1 学科概念的语法分析

学科概念的语法分析是指对学科概念进行词法分析,研究词的组织规律。在学科研究中,学科术语是各个学科为准确表达本领域内的学科概念而创造和使用的专门词语,因此学科概念的语法分析主要体现为对学科术语进行词法分析,即以词素为基本单元,对学科术语的内部构造进行研究。词素是学科术语的构成材料,它们可以单独或联合构成学科术语,绝大多数的学科术语都是由词素联合构成的。此外,术语的内部词素及词素的组合方式决定了其词法结构,进而决定了术语的意义。例如“科学”和“学科”两个术语均由“学”和“科”两个词素构成,但是两个词素的不同组合方式决定了术语意义存在差异。可见,学科概

念表示模型的语法层可以从术语名称、术语成分和术语结构三个维度进行描述,其中术语名称是学科概念的符号表示,术语成分揭示构成学科术语的词素,术语结构描述词素之间的词法结构。

2.2.2 学科概念的语义分析

语义分析的目的是对概念意义进行揭示,以得到学科概念表示模型在语义层的具体描述。概念分析法作为一种意义分析方法,能够通过分解式分析和关联式分析两个方面从内部和外部全面揭示概念的意义^[32]。本文基于概念分析法对学科概念进行语义分析(见图6),以解决当前大多数概念关联方法缺乏学科针对性和语义深度关联性的问题。

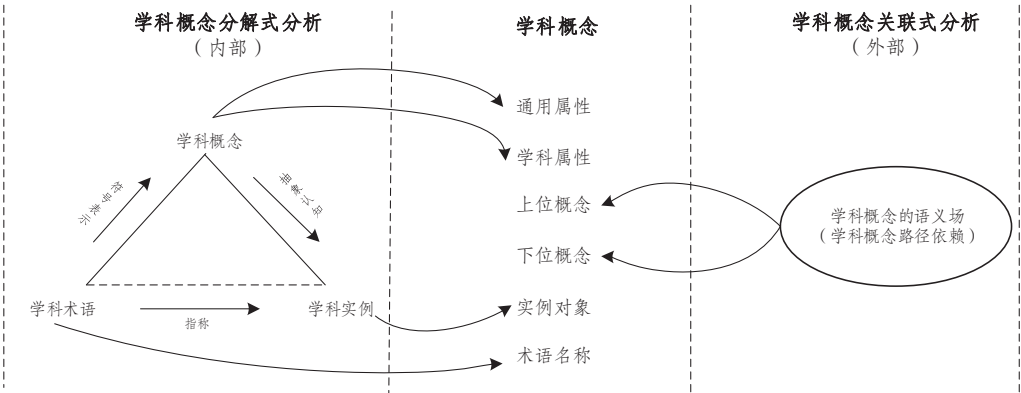


图6 学科概念语义分析

(1) 基于学科语义三角的学科概念分解式分析

Ogden 和 Richards^[33] 提出的语义三角论揭示了概念、符号和对象世界之间的联系,对于学科概念分解式分析具有重要的借鉴意义,将语义三角论迁移应用至具体的学科概念,则可以得到学科语义三角(见图7)。学科语义三角能够揭示出学科术语、学科概念和学科实例之间相互制约、相互作用的关系,包含以下几点含义:①学科概念和学科术语体现了内容和形式的辩证统一关系,学科概念是学科术语表达的思想内容,学科术语是对学科概念的符号表示,二者用实线连接;②学科实例和学科概念体现了物质与意识的辩证统一关系,学科实例是人类认知过程中概念所反映的客观存在,学科概念是对学科实例的抽象认知,是从学科实例中抽象出事物的特有属性而形成的思维形式,二者用实线连接;③学科术语指称学科实例,二者仅仅存在一种间接关系,用虚线连接。这种关系是通过学科研究主体的思维操作过程而发生的,即学科概念是学科实例和学科术语的中介。

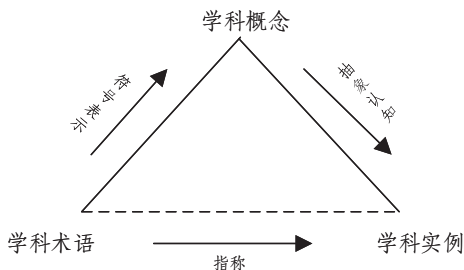


图7 学科语义三角

基于学科语义三角,本文从学科术语、学科概念、学科实例三个维度梳理出学科概念的内部结构。①学科术语揭示学科概念的符号表示,它可与学科概念表示模型的语法层进行融合。②学科概念主要反映概念的内涵,基于属性所反映的学科特征,概念属性可分为通用属性和学科属性两部分。其中,通用属性不拘泥于具体的学科特性,用于描述学科概念的抽象

特征,是学科概念跨学科关联的枢纽。一般而言,学科概念越抽象,普适性越强,跨学科关联的潜力越大。如在艺术世界中,艺术魅力的来源之一就在于其具有的“空筐”性质,所谓“空筐”,就是艺术家为观众提供的想象力弹性空间,是抽象的框架,可以装进世间万物。抽象度高的学科概念好比艺术世界的“空筐”,能够被多个不同的学科所通用,容纳多个学科研究主体的反复科学实践,并根据不同的领域特性去创作升华^[34]。学科属性描述学科概念的具体特征,是概念结合学科特性而形成的属性,能够体现出概念的学科针对性。学科概念的学科属性越显著,概念越具化,越可能成为学科的核心概念。③学科实例是概念所反映的实例对象,与概念内涵共同构成学科概念的内容。由此可见,基于学科语义三角,学科概念内涵在分解式分析下可解析为通用属性、学科属性和实例对象三部分。

(2) 基于语义场理论的学科概念关联式分析

区别于传统的分解式分析,关联式分析旨在通过学科域内概念间的复杂关联来明确学科概念的语义。而语义场理论从系统性角度出发,以研究概念间的相互联系为核心,为此,本文基于语义场理论对学科概念进行学科域内的关联分析。在语义场中,学科概念的语义只有通过它与其他概念的区别和联系才能确定,具体表现为学科概念的路径依赖特性,即学科概念的内涵分析往往依赖于其上位概念和下位概念。具体来说,由于上位概念的概括能力强,是较为抽象的概念,所以在揭示概念时往往是用其上位概念进行解释,以明晰概念的基本范畴;下位概念是对概念的进一步具化,利用下位概念进行概念内涵描述能够揭示概念的实例对象。例如在“动物”这一语义场中,我们在描述概念“狗”时,可以借助上位概念“牲畜”和下位概念“家狗”以明确概念意义。因此,基于语义场理论,学科概念内涵在关联式分析下可表现为学科概念的上位概念和下位概念。

综上,基于学科语义三角的学科概念分解

式分析和基于语义场理论的学科概念关联式分解的讨论结果,本文得到了学科概念表示模型在语义层面的具体描述内容,即通用属性、学科属性、学科实例、上位概念和下位概念。

2.2.3 学科概念的语用分析

学科概念的语用分析是对学科概念的使用情境进行分析,本文通过学科概念的语用分析来得到学科概念表示模型在语用层面的具体描述内容。在跨学科概念关联中发现,学科概念的语用情境是对概念学科特征的进一步明确,能够体现出学科概念跨学科关联时的学科针对性。具体来说,学科概念的使用情境包括研究对象情境和研究主题情境。一方面,学科研究对象的特殊性是一门学科区别于其他学科的首要条件,在学科概念的跨学科关联分析中,研究对象情境主要是明确学科概念所属的具体学科;另一方面,研究主题是在某一点、某一面 on 对学科研究对象展开研究,是对学科研究对象的进一步收敛,研究主题情境主要是明确学科概念所属的研究主题。因此在学科概念表示模型中,语用层也相应表现为研究对象和研究主题,二者分别揭示了学科概念所属具体学科和具体主题。

2.2.4 学科概念表示模型

在上述研究的基础上,本文从语法—语义—语用三个层面构建了学科概念表示模型(见表1)。具体来说,语法层揭示学科概念的形式结构,包括术语名称、术语成分和术语结构;语义层描述学科概念在学科背景下的内涵意义,包括通用属性、学科属性、学科实例、上位概念和下位概念;语用层揭示了学科概念的使用情境,包括研究对象和研究主题。综合来看,学科概念表示模型中的语义层和语用层充分反映了学科概念的学科特征,具有学科针对性的显著优势,而语义层又从概念内部结构和学科域内概念外部关联的角度实现了概念语义的深度揭示。因此,本文提出的学科概念表示模型能够多维度、深层次和有针对性地描述学科概念,进而可以为学科概念跨学科关联提供有力支撑。

表1 学科概念表示模型

结构层次	描述内容
语法层	术语名称::= 学科概念的符号表示 术语成分::= 学科术语的构成词素 术语结构::= 词素之间的词法结构
语义层	通用属性::= 学科概念的抽象属性 学科属性::= 学科概念的具体属性 学科实例::= 学科概念的实例对象 上位概念::= 学科概念的属概念 下位概念::= 学科概念的种概念
语用层	研究对象::= 学科概念所属具体学科 研究主题::= 学科概念所属具体主题

2.3 学科概念跨学科关联方法

基于学科概念表示模型,本文提出了学科概念跨学科关联方法,具体包括语法、语义和语用三个层面的跨学科关联计算。

2.3.1 语法层的跨学科关联计算

学科概念表示模型的语法层描述了概念的术语名称、术语成分和术语结构,所以跨学科概念对(由来自不同学科的两个概念组成)在语法上的关联计算应以词素为单位,在考虑术语词法结构的前提下进行语法相似度计算。跨学科概念对的语法相似度计算建立在公共词素的基础上^[35],它受公共词素的数量及其权重两个因素的影响。具体而言,跨学科概念对 (x, y) 的语法相似度计算见公式(1),其中 A, B 分别为概念 x, y 的词素集合, C 为两个概念的公共词素, S 为集合中词素的权重之和,例如 S_A 表示 A 中的词素权重之和,以此类推。

$$S_{\text{语法}}(x, y) = 2S_C / (S_A + S_B) \quad (1)$$

词素的权重由术语结构决定,术语结构中不同词素对术语的整体表达作用不同,权重也不同。术语包括单词型和词组型两种类型,为基于术语结构识别词素权重,本文将单词型和词组型术语结构归纳为前焦型、后焦型和并焦型三种类型(见表2)。在前焦型术语结构中,前一部分的权重比后一部分的权重大;在后焦型术语结构中,后一部分的权重比前一部分的权重大;并焦型术语的各个部分权重相等。

表2 术语结构分类

单词型术语结构 ^[36]	词组型术语结构 ^[37]	表示关系	术语结构归纳 ^[38]
支配式	述宾结构	支配关系	前焦型
补充式	述补结构	补充关系	
陈述式	主谓结构	陈述关系	后焦型
偏正式	偏正结构	修饰关系	
并列式	联合结构	并列关系	并焦型

为计算词素权重,本文采用二叉树来描述术语结构。具体而言,二叉树的顶层节点为术语,之后每一层都采用二分的方法对术语进行切分,直至左右子树的底层节点全部被切分为词素为止。基于二叉树,词素权重分配规则为:①首先确定左子树底层的最左侧节点权重为2;②对任意节点而言(根节点除外),与左侧节点为前焦型结构的节点,其权重为左侧节点权重-1;与左侧节点为后焦型结构的节点,其权重为左侧节点权重+1;与左侧节点为并焦型结构的节点,其权重与左侧节点保持一致;③父节点的权重为子节点权重求和。

2.3.2 语义层的跨学科关联计算

依据学科概念表示模型的语义层内容,跨学科概念对在语义上的比较包含属性、实例和结构相似度计算三个部分。

(1) 跨学科概念对的属性相似度计算

在学科概念表示模型的语义层中属性区分为通用属性和学科属性,通用属性较为抽象,学科属性较为具体,而属性越具体对概念的相似度

计算贡献度越大,因此在概念属性相似度计算时要考虑属性在跨学科关联中的不同重要性。

① 属性重要性程度计算

本文借鉴 DR-DC^[39] 算法思想计算出属性的学科归属度,以此区分出概念的通用属性和学科属性的相对重要性程度。学科归属度较高的属性具有两个特征:第一,该属性在本学科领域出现的频率高于在其他学科领域出现的频率(也即领域相关性);第二,该属性在领域内被广泛使用(也即领域一致性)。假设存在 n 个学科 $\{D_1, D_2, \dots, D_n\}$, 则属性 a 与学科 D_i 的领域相关度计算方式见公式(2)。其中 $P(a|D_i)$ 表示属性 a 在学科领域 D_i 中出现的频率。

$$DR(a, D_i) = P(a|D_i) / \max_{1 \leq j \leq n} P(a|D_j) \quad (2)$$

假设学科 D_i 中的概念集合为 $\{c_1, c_2, \dots, c_m\}$, 则属性 a 与学科 D_i 的领域一致度的计算见公式(3)。其中 c_j 表示从 D_i 中随机抽取的一个概念, $P(a, c_j)$ 表示属性 a 在概念 c_j 中出现的频率。

$$DC(a, D_i) = H(P(a, c_j)) = \sum_{c_j \in D_i} (P(a, c_j) \log \frac{1}{P(a, c_j)}) \quad (3)$$

假设具有属性 a 的学科集合为 $\{D_1, D_2, \dots, D_r\}$, 则属性 a 的重要性程度计算见公式(4)。

$$I_a = \sum_{i=1}^r (DR(a, D_i) + DC(a, D_i)) / r \quad (4)$$

② 属性相似度计算

跨学科概念对具有的共有属性越多,且共有属性的权重越大,属性相似度越大。跨学科概念对 (x, y) 的属性相似度计算见公式(5)。其中, $f(x \cap y)$ 为 x, y 的共有属性个数, $f(x - y)$ 为 x 有但 y 没有的属性个数, $f(y - x)$ 则相反, A, B 分别为概念 x 和 y 的所属学科。

$$S_a(x, y) = \frac{\sum_{i=1}^{f(x \cap y)} I_i}{\sum_{i=1}^{f(x \cap y)} I_i + \sum_{j=1}^{f(x-y)} S(j, D_A) + \sum_{k=1}^{f(y-x)} S(k, D_B)} \quad (5)$$

(2) 跨学科概念对的实例相似度计算

本文采用 Glue 系统^[40]提出的联合分布概率方法来计算实例相似度,见公式(6),其中 X 、 Y 分别为概念 x 和 y 的实例集合, $P(X, Y)$ 表示所有实

例中既属于 X 又属于 Y 的概率, $P(X, \bar{Y})$ 表示所有实例中属于 X 但不属于 Y 的概率, $P(\bar{X}, Y)$ 表示所有实例中属于 Y 但不属于 X 的概率。

$$S_i(x, y) = P(X \cap Y) / P(X \cup Y) = P(X, Y) / (P(X, Y) + P(X, \bar{Y}) + P(\bar{X}, Y)) \quad (6)$$

(3) 跨学科概念对的结构相似度计算

跨学科概念对的结构相似度计算分为上位概念和下位概念的相似度计算两部分,概念 x 和

概念 y 结构相似度($S_s(x, y)$)由概念的上位概念和下位概念共同决定,计算见公式(7)。

$$S_s(x, y) = \frac{c}{|I(x)| |I(y)|} \sum_{i=1}^{|I(x)|} \sum_{j=1}^{|I(y)|} S(I_i(x), I_j(y)) + \frac{c}{|O(x)| |O(y)|} \sum_{i=1}^{|O(x)|} \sum_{j=1}^{|O(y)|} S(O_i(x), O_j(y)) \quad (7)$$

其中 $|I(x)|$ 和 $|I(y)|$ 分别表示概念 x 和概念 y 上位概念的个数, $I_i(x)$ 和 $I_j(y)$ 分别表示概念 x 与概念 y 对应的上位概念。 $S(I_i(x), I_j(y))$ 表示两上位概念的相似度, $|O(x)|$ 和 $|O(y)|$ 分别表示概念 x 和概念 y 下

位概念的个数, $O_i(x)$ 和 $O_j(y)$ 分别表示概念 x 与概念 y 对应的下位概念。 $S(O_i(x), O_j(y))$ 表示两下位概念的相似度, c 为退变系数。

综上,跨学科概念对 (x, y) 的综合语义相似度计算见公式(8),其中 $w_1 + w_2 + w_3 = 1$ 。

$$S_{yti}(x, y) = w_1 S_a(x, y) + w_2 S_i(x, y) + w_3 S_s(x, y) \quad (8)$$

2.3.3 语用层的跨学科关联计算

由前文可知,跨学科概念对在语用上的关联通过比较概念所属的研究主题来实现,而学科概念往往从属于多个研究主题,可用主题词集合来表示,因此本文将跨学科概念对在语用上的相似度计算转化为两个主题词集合间的相似度衡量。

(1) 相同主题词的相似度计算

两个主题词集合共有的主题词越多,集合相似度越大,本文采用 Jaccard 系数^[7]计算相同主题词的相似度,如公式(9)所示,其中 T_x 、 T_y 分别为概念 x 和 y 的主题词集合。

$$S_{st}(x, y) = |T_x \cap T_y| / |T_x \cup T_y| \quad (9)$$

(2) 相异主题词的相似度计算

相异主题词之间可能存在潜在的语义关联,概念 x 和概念 y 在相异主题词上的相似度为两个主题词集合中两两相异主题词相似度的均值,其中两相异主题词间的关联度由二者间的互信息来衡量,具体计算见公式(10),其中 $P(t_i(x), t_j(y))$ 为相异主题词 $t_i(x)$ 和 $t_j(y)$ 共同出现的文档在整个文档集合中出现的概率, $P(t_i(x))$ 和 $P(t_j(y))$ 分别为 $t_i(x)$ 和 $t_j(y)$ 出现的文档在整个文档集合中出现的概率, k 和 l 分别为两个主题词集合包含的相异主题词的个数。

$$S_{sd}(x, y) = \frac{\sum_{i=1}^k \sum_{j=1}^l \log(P(t_i(x), t_j(y)) / (P(t_i(x)) P(t_j(y))))}{kl} \quad (10)$$

基于上述分析,跨学科概念对在研究主题 上的相似度计算见公式(11)。

$$S_{yyong}(x, y) = S_{st}(x, y) + S_{sd}(x, y) \quad (11)$$

2.3.4 学科概念跨学科关联度计算及其关联方式判定

(1) 学科概念跨学科关联度计算

通过对跨学科概念对的语法、语义和语用

$$S(x, y) = w_{yfa} S_{yfa}(x, y) + w_{yyi} S_{yyi}(x, y) + w_{yyong} S_{yyong}(x, y) \quad (12)$$

(2) 学科概念跨学科关联方式判定

根据跨学科概念对在语法层、语义层和语用层上是否相关,可以把学科概念的跨学科关联方式划分为多种类型。根据学科概念在语法层是否关联分为显性关联和隐性关联,根据

上的比较,可以计算出跨学科概念对的跨学科关联度,具体计算见公式(12),从而识别出跨学科概念关联关系。其中, $w_{yfa} + w_{yyi} + w_{yyong} = 1$ 。

学科概念在语义层是否关联分为本质关联和非本质关联,根据学科概念在语用层是否关联分为同域关联和异域关联,本文归纳出学科概念跨学科关联的七种方式(见图8)。

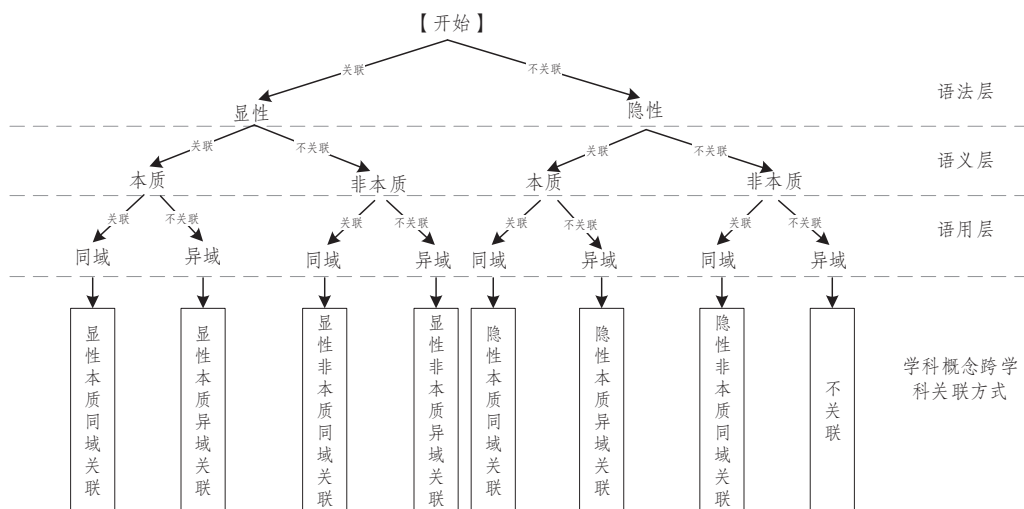


图8 学科概念的跨学科关联方式

①显性本质同域关联是指不同学科概念在语法层、语义层和语用层上都具有关联关系,是学科概念跨学科关联最为紧密的一种方式;②显性本质异域关联是指不同学科概念在语法层和语义层上相关,在语用层不相关,这主要是由于学科概念所属的学科研究主题差异较大;③显性非本质同域关联是指不同学科概念在语法层和语用层上相关,在语义层不相关,这是由于不同的学科概念本身在内涵意义上不相关,

但是从属于相同的研究主题;④显性非本质异域关联是指不同学科概念仅在语法层相关,在语义层和语用层不相关,此种跨学科概念关联方式较为浅显;⑤隐性本质同域关联是指不同学科概念在语法层不相关,而在语义层和语用层相关,发现这种概念关联有利于学科概念的跨域交流;⑥隐性本质异域关联是指不同学科概念在语法层和语用层不相关,在语义层相关,这主要是由于不同学科概念对相似的实例对象

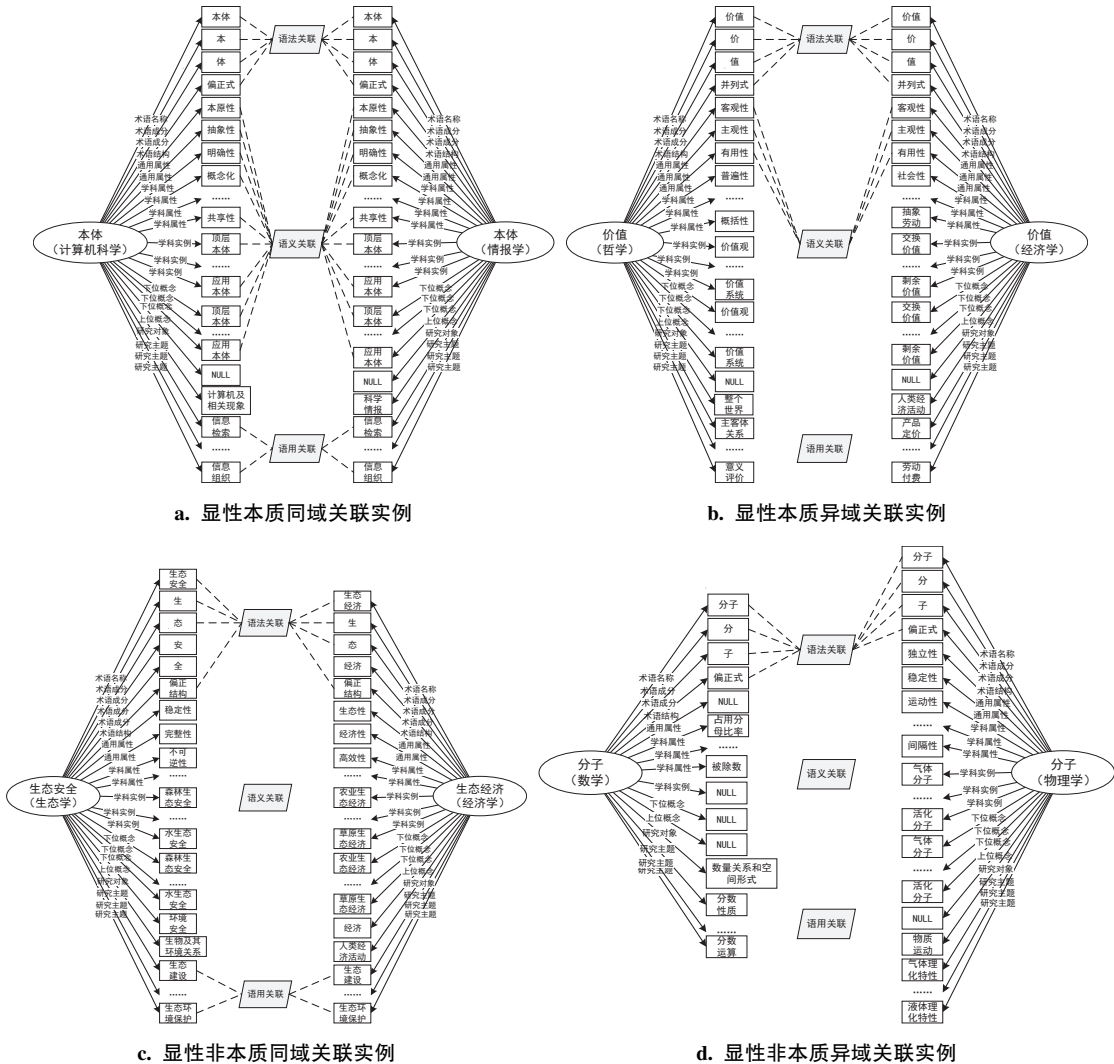
进行各自归纳;⑦隐性非本质同域关联是指不同学科概念仅在语用层相关,这是由于不同学科概念对同一研究主题的不同方面进行研究。

3 应用案例及讨论

3.1 应用案例

为验证本文提出的学科概念表示模型、学科概念跨学科关联方法和学科概念跨学科关联方式的适用性和科学性,笔者从我国学科体系中,随机抽取情报学—计算机科学、哲学—经济

学、数学—物理学、生态学—经济学、数学—计算机科学、信息科学—建筑科学等学科对,进行跨学科关联分析。实证过程包括:①确定各个学科概念在语法层、语义层和语用层的具体内容,构建学科概念表示模型;②基于学科概念表示模型,依据学科概念跨学科关联方法计算跨学科概念对在语法、语义和语用三个层面上的相关性;③根据跨学科概念对在三个层面的相关性,判定学科概念的跨学科关联方式。最后,本文从识别出的跨学科概念对中选出七组典型的跨学科概念对作为案例进行分析(见图9)。



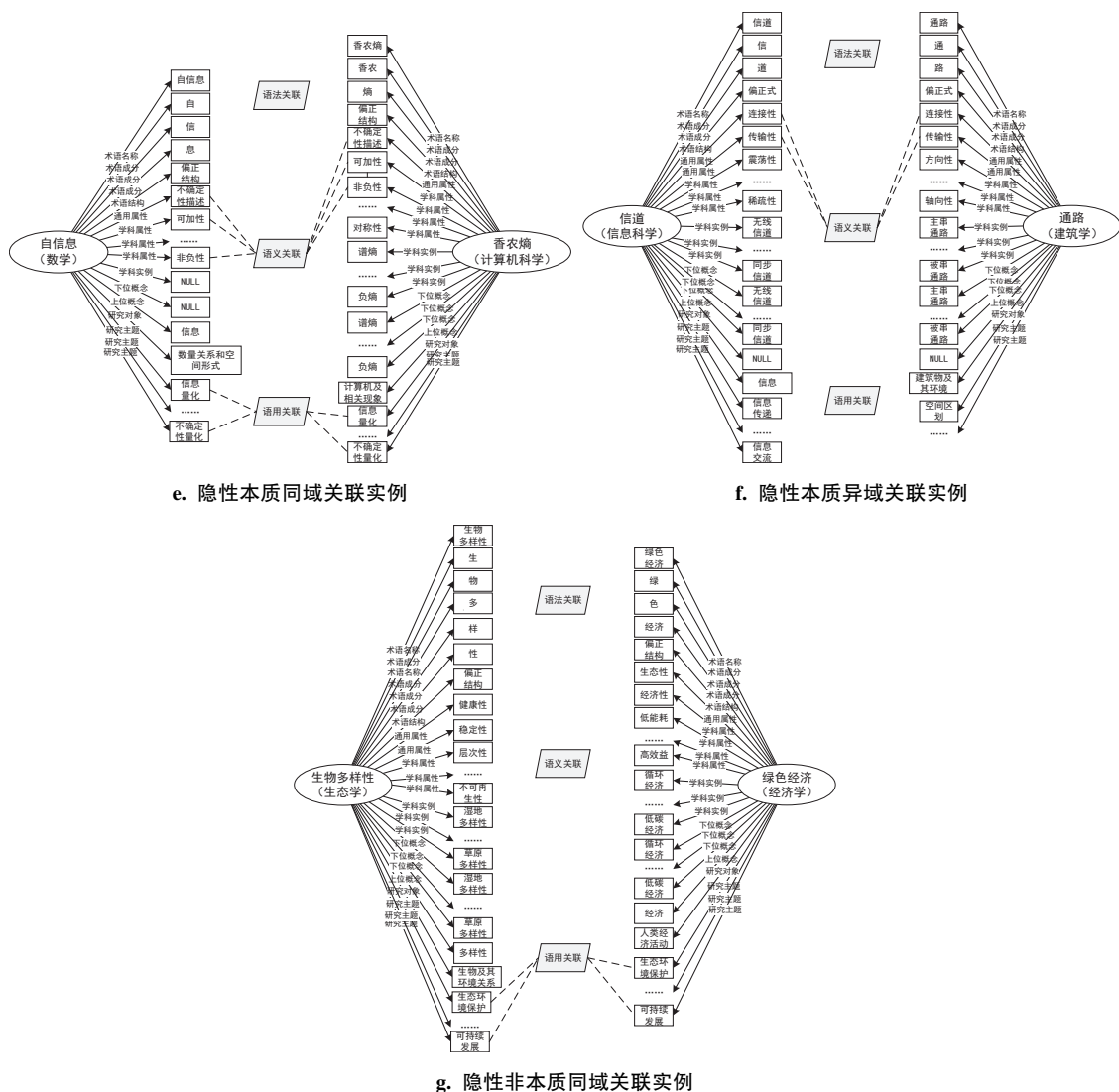


图9 跨学科概念关联实证结果

在图9的子图a中,计算机科学“本体”概念和情报学“本体”概念在语法、语义和语用上显现出非常紧密的相关性,因此属于跨学科的显性本质同域关联,这主要是由于情报学和计算机科学在信息组织、信息检索等方面具有相同的研究主题,因而情报学将计算机科学“本体”概念移植至本学科以进行相同主题的研究。

在子图b中,哲学“价值”概念和经济学“价值”概念在语法和通用属性上关联,属于跨学科

的显性本质异域关联,这是由于经济学“价值”概念是由哲学“价值”概念跨学科演绎产生的,在学科概念跨学科演绎过程中,哲学“价值”概念的通用属性被保留的同时,结合经济学的学科特征形成了经济学的“价值”概念。

在子图c中,生态学“生态安全”概念和经济学“生态经济”概念在语法上具有公共的词素和词法结构,在语义上不存在任何关联关系,在语用上都从属于“生态建设”“生态环境保护”等

研究主题,因而属于跨学科显性非本质同域关联。

在子图 d 中,数学“分子”概念和物理学“分子”概念只在语法上具有关联关系,属于跨学科显性非本质异域关联,由此可知在不同学科中,名称相同的概念其意义不一定相同。

在子图 e 中,数学“自信息”概念和计算机科学“香农熵”概念在语法上没有公共的词素,不存在语法上的关联,在语义上具有公共的通用属性和学科属性,在语用上都从属于“信息量化”等研究主题,因此属于跨学科隐性本质同域关联。

在子图 f 中,信息科学“信道”概念和建筑学“通路”概念仅在通用属性上存在关联关系,属于跨学科隐性本质异域关联,这主要是由概念归纳导致的跨学科概念关联,即信息科学和建筑学在研究空间路径这一实例对象上具有关联性,因此两个学科对相似的实例对象进行抽象认识所形成的概念在属性上存在跨学科关联。

在子图 g 中,生态学“生物多样性”概念和经济学“绿色经济”概念仅在语用上关联,属于跨学科隐性非本质同域关联,这主要是由于“生态环境保护”“可持续发展”等研究主题的开展涉及多个学科的知识,因而不同学科对同一研究主题的不同侧面进行研究形成了不同的学科概念。

3.2 讨论

通过应用案例分析可以看出,本文在对学科概念跨学科关联形成机制进行深入分析的基础上提出的学科概念表示模型、学科概念跨学科关联方法和学科概念跨学科关联方式,能够帮助人们深入理解和有效实现不同学科概念的跨学科关联,具有一定优势。

(1)跨学科性。为真正打破学科藩篱,建立学科概念间的跨学科联系,本文从学科概念归纳和学科概念移植两种学科概念生成方式出发,分析了学科概念跨学科关联的形成机制。在此基础上,指出学科概念跨学科关联主要表

现在学科概念名称、属性、外延和情境的关联上,也即概念的语法、语义(包括属性、外延)和语用三个层面的关联,这为真正实现学科概念的跨学科关联指明了路径,有利于学科概念跨学科关联的顺利实现。

(2)多维性。当前的概念关联方法大多是从语法、语义等单一维度进行分析,造成多维度概念关联分析能力的缺失,影响了概念关联的科学性。本文在分析学科概念跨学科关联形成机制的基础上,从语法—语义—语用三个层面构建学科概念跨学科关联中的学科概念表示模型,以实现学科概念多维度的全息表达。其中,语法层主要揭示学科概念的形式结构,语义层描述学科概念在学科背景下的内涵意义,语用层主要揭示学科概念的使用情境。该概念表示模型能够在一定程度上实现多角度、全方位的学科概念跨学科关联。

(3)学科针对性。现有的概念关联方法大多不具有学科针对性,发现的概念关联不能完全体现出概念的学科特性。为解决这一问题,本文在语义层从通用属性、学科属性、学科实例、上位概念和下位概念五个方面对概念在具体学科中的语义内涵进行全方位揭示,在语用层从研究对象和研究主题两个方面对学科概念的学科情境进行明确。在此基础上,对学科概念跨学科关联进行判断,并充分结合了学科概念的学科特征,具有学科针对性,能够实现学科概念跨学科关联关系的有效发现。

(4)语义性。在当前的概念关联研究中,语义关联是主流方向,但是大都没有实现概念语义的深度关联,因而很难满足学科概念跨学科关联的需求。本文基于概念分析方法,分别从分解式分析和关联式分析两个方面探讨学科概念的内部结构和学科域内的外部关联,以此进行学科概念语义的深度挖掘,进而从语义深度关联上支撑学科概念的跨学科关联。

(5)系统性。本文根据跨学科概念在语法层、语义层和语用层上是否相关,归纳出学科概念跨学科关联方式的所有类型;显性本质同域

关联、显性本质异域关联、显性非本质同域关联、显性非本质异域关联、隐性本质同域关联、隐性本质异域关联和隐性非本质同域关联,这有利于系统实现学科概念之间各种类型的跨学科关联。

4 结论与展望

现如今,跨学科研究是综合性解决人类面临的复杂问题的重要手段,是引发创新的重要源泉,已成为科学发展的重要趋势。在这一背景下,人们对跨学科知识服务的需求越来越迫切,急需图书情报人员帮助他们解决跨学科知识的存取问题。这就需要发现不同学科概念间的关联,构建跨学科知识组织系统,进行跨学科的知识组织。为此,本文对学科概念跨学科关联问题进行了全方位、系统化和深层次的探索。首先,从概念归纳和概念移植两种概念生成方式出发,分析了学科概念跨学科关联的形成机制;而后,从语法—语义—语用角度构建了学科概念跨学科关联中的学科概念表示模型,以实现学科概念的全息表达;接着,基于学科概念表示模型提出了学科概念跨学科关联的方法,

概括学科概念跨学科关联的七种方式;最后,通过应用实例检验了概念表示模型和概念关联方法的适用性和科学性。

学科概念跨学科关联的探讨只是跨学科知识组织研究的开端,在今后的研究中还需进一步探讨以下两个方面的问题。①跨学科知识组织系统中常识知识的嵌入问题。现有的知识组织方法难以支撑跨学科知识服务的一个重要原因就在于只重视明显的学科专业知识,而忽视了基础的学科常识知识,导致跨学科知识存取存在困难,原因在于一个为某学科领域中专家所熟知的常识,其他学科领域的专家并不一定知道。因此,如何获取学科领域中的常识知识,并将它们有机嵌入到跨学科知识组织系统之中,成为跨学科知识组织研究需要进一步解决的关键问题。②跨学科知识组织系统的动态更新和维护问题。随着科学的发展,各学科不断发现新的知识,新的学科概念和关联不断涌现,若要有效实现对学科知识的跨学科组织,就需要解决跨学科知识组织系统的动态更新和维护问题,从而有效支撑跨学科知识组织,促进跨学科的知识深度融合,开展跨学科知识服务,进而推进跨学科研究的深入发展。

参考文献

- [1] 马文峰,杜小勇.关于知识组织体系的若干理论问题[J].中国图书馆学报,2007(2):13-17,46. (Ma W F, Du X Y. Some theoretical issues relating to knowledge organization system[J]. Journal of Library Science in China, 2007(2):13-17,46.)
- [2] 文庭孝,刘晓英,刘进军.知识关联的理论基础研究[J].图书馆,2010(4):9-11. (Wen T X, Liu X Y, Liu J J. Research on theory foundation of knowledge connection[J]. Library, 2010(4):9-11.)
- [3] Rokaya M, Atlam E, Fuketa M, et al. Ranking of field association terms using co-word analysis[J]. Information Processing & Management, 2008, 44(2):738-755.
- [4] 王立学,孙杨,杨代庆.基于引文的情报学领域主题关联特征分析[J].情报杂志,2012,31(10):27-31. (Wang L X, Sun Y, Yang D Q. Capturing topics linkage of information science based on citations[J]. Journal of Intelligence, 2012, 31(10):27-31.)
- [5] 叶鹰,张力,赵星.用共关键词网络揭示领域知识结构的实验研究[J].情报学报,2012,31(12):1245-1251. (Ye Y, Zhang L, Zhao X. An experimental study on revealing domain knowledge structure by co-keyword networks[J]. Journal of the China Society for Scientific and Technical Information, 2012, 31(12):1245-1251.)
- [6] Levenshetin V I. Binary codes capable of correcting deletions, insertions, and reversals[J]. Soviet Physics Doklady, 1966, 10(8):707-710.

- [7] Jaccard P. The distribution of the flora in the alpine zone[J]. New Phytologist, 1912, 11(2): 37-50.
- [8] 刘萍, 陈烨. 词汇相似度研究进展综述[J]. 现代图书情报技术, 2012(Z1): 82-89. (Liu P, Chen Y. Survey of the state of the art in word similarity[J]. New Technology of Library and Information Service, 2012(Z1): 82-89.)
- [9] 王瑞琴, 杨小明, 楼俊钢. 词汇语义相关性度量研究[J]. 情报学报, 2016, 35(4): 389-404. (Wang R Q, Yang X M, Lou J G. Research of lexical semantic relatedness measurement[J]. Journal of the China Society for Scientific and Technical Information, 2016, 35(4): 389-404.)
- [10] 韩普, 王东波, 王子敏. 词汇相似度计算和相似词挖掘研究进展[J]. 情报科学, 2016, 34(9): 161-165. (Han P, Wang D B, Wang Z M. Research advancement in word similarity calculation and mining[J]. Information Science, 2016, 34(9): 161-165.)
- [11] 李慧. 词语相似度算法研究综述[J]. 现代情报, 2015, 35(4): 172-177. (Li H. A review on the research of word similarity algorithms[J]. Journal of Modern Information, 2015, 35(4): 172-177.)
- [12] Salton G. A vector space model for automatic indexing[J]. Communications of the ACM, 1975, 18(11): 613-620.
- [13] Deerwester S, Dumais S T, Furnas G W, et al. Indexing by latent semantic analysis[J]. Journal of the Association for Information Science & Technology, 2010, 41(6): 391-407.
- [14] 韩普, 王东波, 朱恒民. 基于复杂网络的汉语相似词挖掘和相似度计算研究[J]. 情报学报, 2015, 34(8): 885-896. (Han P, Wang D B, Zhu H M. Research of Chinese similar words mining and similarity calculation based on complex network[J]. Journal of the China Society for Scientific and Technical Information, 2015, 34(8): 885-896.)
- [15] Rada R, Mili H. Development and application of a metric on semantic nets[J]. IEEE Transactions on Systems, Man, and Cybernetics, 1989, 19(1): 17-30.
- [16] Tversky A. Features of similarity[J]. Psychological Review, 1977, 84(4): 327-352.
- [17] Resnik P. Using information content to evaluate semantic similarity in a taxonomy[C]//Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence. 1995: 448-453.
- [18] Rodriguez M A. Assessing semantic similarity among spatial entity classes[D]. Maine: University of Maine, 2000.
- [19] Li Y, Bandar Z A, McLean D. An approach for measuring semantic similarity between words using multiple information sources[J]. IEEE Transactions on Knowledge and Data Engineering, 2003, 15(4): 871-882.
- [20] 张硕望, 欧阳纯萍, 阳小华, 等. 融合知网和搜索引擎的词汇语义相似度计算[J]. 计算机应用, 2017, 37(4): 1056-1060. (Zhang S W, Ouyang C P, Yang X H, et al. Word semantic similarity computation based on integrating HowNet and search engines[J]. Journal of Computer Applications, 2017, 37(4): 1056-1060.)
- [21] 吴克介, 王家伟. 基于知网与搜索引擎的词汇语义相似度计算[J]. 计算机与现代化, 2018(4): 90-94. (Wu K J, Wang J W. Vocabulary semantic similarity computation based on HowNet and search engine[J]. Computer and Modernization, 2018(4): 90-94.)
- [22] Swanson D R. Undiscovered public knowledge[J]. Library Quarterly, 1986, 56(2): 103-118.
- [23] 荣毅虹, 梁战平. 基于文献的发现[J]. 情报学报, 2002(4): 386-390. (Rong Y H, Liang Z P. Literature-based discovery[J]. Journal of the China Society for Scientific and Technical Information, 2002(4): 386-390.)
- [24] 马明, 武夷山. Don R. Swanson 的情报学学术成就的方法论意义与启示[J]. 情报学报, 2003, 22(3): 259-266. (Ma M, Wu Y S. The methodological enlightenment and significance of Don R. Swanson's achievements in information science[J]. Journal of the China Society for Scientific and Technical Information, 2003, 22(3): 259-266.)
- [25] Swanson D R. Fish oil, Raynaud's Syndrome, and undiscovered public knowledge[J]. Perspectives in Biology & Medicine, 1986, 30(1): 7-18.
- [26] Swanson D R. Migraine and magnesium: eleven neglected connections[J]. Perspectives in Biology & Medicine, 1988, 31(4): 526-557.

- [27] 黄水清,熊健,李志燕. 闭合式非相关文献知识发现方法在中文文献中的验证[J]. 中国图书馆学报,2007(5):83-87. (Huang S Q,Xiong J,Li Z Y. Verification of closed non-interactive literature knowledge discovery method in Chinese literature[J]. Journal of Library Science in China,2007(5):83-87.)
- [28] 黄水清,程冲,李志燕. 开放式非相关文献知识发现方法在中文文献中的验证[J]. 情报理论与实践,2008(2):246-250. (Huang S Q,Cheng C,Li Z Y. Application of the open knowledge discovery method for non-inter-related literature in Chinese literature[J]. Information Studies:Theory & Application,2008(2):246-250.)
- [29] 刘小慧,李长玲,崔斌,等. 基于闭合式非相关知识发现的潜在跨学科合作研究主题识别——以情报学与计算机科学为例[J]. 情报理论与实践,2017,40(9):71-76. (Liu X H,Li C L,Cui B,et al. Research topics identification of potential interdisciplinary collaboration based on closed and irrelevant knowledge discovery[J]. Information Studies:Theory & Application,2017,40(9):71-76.)
- [30] 张云秋,冷伏海. 非相关文献知识发现的理论基础研究[J]. 中国图书馆学报,2009,35(4):25-30. (Zhang Y Q,Leng F H. A study on theoretical foundation of noninteractive literature-based discovery[J]. Journal of Library Science in China,2009,35(4):25-30.)
- [31] 李勇,冷伏海,王林. 基于非相关文献的三阶知识发现方法探讨[J]. 中国图书馆学报,2011,37(4):21-26,69. (Li Y,Leng F H,Wang L. A study on third-order disjoint literature-based discovery method[J]. Journal of Library Science in China,2011,37(4):21-26,69.)
- [32] 黄远帆. 论两类概念分析:关联式与分解式[J]. 河北学刊,2015,35(2):29-36. (Huang Y F. On two forms of conceptual analysis:connective analysis and decompositional analysis[J]. Hebei Academic Journal,2015,35(2):29-36.)
- [33] Ogden C K,Richards I A. The meaning of meaning;a study of the influence of language upon thought and of the science of symbolism[M]. New York:Harcourt,Brace & World,Inc. ,1923:1-23.
- [34] 曾月新. 熵概念的跨学科发展[J]. 天津师大学报(社会科学版),1995(1):42-45. (Zeng Y X. Interdisciplinary development of the concept of entropy[J]. Journal of Tianjin Normal University(Social Sciences),1995(1):42-45.)
- [35] 朱毅华. 智能搜索引擎中的同义词识别算法研究[D]. 南京:南京农业大学,2001. (Zhu Y H. Automatic recognition of synonym in construction of intelligent search engine[D]. Nanjing:Nanjing Agricultural University,2001.)
- [36] 冯志伟. 汉语单词型术语的结构[J]. 科技术语研究,2004(1):15-20. (Feng Z W. The structure of Chinese word terms[J]. China Terminology,2004(1):15-20.)
- [37] 冯志伟. 汉语词组型术语的结构[J]. 科技术语研究,2004(2):35-37. (Feng Z W. The structure of Chinese phrase terms[J]. China Terminology,2004(2):35-37.)
- [38] 冯志伟. 中文科技术语描述中的三种结构[J]. 语文建设,1989(5):14-20. (Feng Z W. Three structures in the description of Chinese scientific and technological terms[J]. Language Planning,1989(5):14-20.)
- [39] Basili R, Missikoff M, Velardi P. Identification of relevant terms to support the construction of domain ontologies [C]//ACL Workshop on HLT,Toulouse,France. 2001.
- [40] Doan A H, Madhavan J, Domingos P, et al. Learning to map between ontologies on the semantic web[C]//Proceedings of the 11th International Conference on World Wide Web. 2002:662-673.

王忠义 华中师范大学信息管理学院副教授,硕士生导师。湖北 武汉 430079。

彭思源 华中师范大学信息管理学院硕士研究生。湖北 武汉 430079。

夏立新 华中师范大学信息管理学院教授,博士生导师。湖北 武汉 430079。

(收稿日期:2021-10-12)