第43卷 第5期
2016年5月
计 算 机 科 学
Computer Science
Vol. 43 No. 5
May 2016

基于范畴论的本体描述方法

余珊珊[1] 苏锦钿[2] 易法令[1]

(广东药学院医药信息工程学院 广州 510006)[1] (华南理工大学计算机科学与工程学院 广州 510640)[2]

摘 要 针对各种本体语言在描述本体过程中容易产生理解不一致的问题,利用范畴论的抽象性及图形化表示的优势,提出一种基于范畴论的本体描述方法,并给出了本体、本体映射和本体实例化的范畴论定义,其中本体被描述成范畴中的对象,本体间的映射被描述成范畴中对象间的同态射,本体实例化被描述成范畴间的函子。在此基础上,结合共极限和推出给出本体合并的描述及解释,并证明了本体合并的一些典型性质。

关键词 本体,范畴论,本体描述,本体合并,本体对齐

中图法分类号 TP301.2 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.5.007

Descriptions for Ontologies Based on Category Theory

YU Shan-shan[1] SU Jin-dian[2] YI Fa-ling[1]

(College of Medical Information Engineering, Guangdong Pharmaceutical University, Guangzhou 510006, China)[1]

(College of Computer Science and Engineering, South China University of Technology, Guangzhou 510640, China)[2]

Abstract To solve the inconsistence problems brought by various ontology languages when describing ontologies, a description method for ontologies based on category theory was proposed by taking full advantages of abstractness and graphical expressions of category theory. The category theoretical definitions for ontology, ontology mapping and ontology instanilization were presented, where each ontology is described as an object of a category, each morphism between ontologies is described as the a homomorphism between objects of a category, and the instanlization of ontologies is described as a functor between two categories. After that, the notions of colimit and pushout were used to describe the merging of ontologies. And some interesting properties for ontology merging were proved from the viewpoints of category theory.

Keywords Ontology, Category theory, Ontology description, Ontology merging, Ontology alignment

1 引言

目前针对本体的描述方法主要分为逻辑方法和代数方法,其数学理论基础本质上都可归结为集合论。由于集合论本身需要依赖于集合的一些特殊性质,对某些问题域语义的表述较为复杂晦涩,因此一方面降低了本体的重用性,另一方面缺乏直观性,影响了本体的集成。作为近年来数学领域的一个新兴研究方向,范畴论(Category Theory)[1]被公认为可以取代集合论并成为计算机科学中新的数学理论基础。它不仅可以构成集合论、代数和共代数等的数学理论基础,且具有更高的抽象性和更强、更直观的描述能力。范畴论的研究重点是对象及对象间的关系,而非对象的内部结构。这意味着范畴论可以作为一种通用的元规范语言用于抽象地描述本体及本体间的各种关系,并且独立于各种本体描述语言及具体的实现[2],从而为本体集成提供一个统一和抽象的数学框架及语义模型;而且,范畴论本身非常适合于从图形化的角度对本体间的各种关系及性质进行推理和验证。

早期范畴论主要应用于程序语言及类型理论,例如文献[3-7]。近年来,一些学者开始将范畴论应用于本体的相关研究。例如,文献[8]指出范畴论可作为本体形式化描述的一种合适的理论基础。文献[9,10]初步探讨了范畴论在本体研究中的应用,并与集合论进行比较,指出范畴论在本体研究中的优势。这些工作初步探讨了范畴论应用于本体研究的可行性,没有详细地给出本体中各主要概念的范畴论解释。在本体映射研究方面,文献[11]利用代数规范间的射分析了本体间的关系,并结合范畴论中的 span 给出本体对齐的定义。文献[12]在文献[11]的基础上利用“桥接公理”(bridge axioms)将本体间的 V-对齐扩展为 W-对齐,并给出一些复杂对齐关系的描述。文献[13]则对各种本体映射关系间的兼容性和不兼容性进行分析。在本体合并研究方面,文献[2,14-16]结合拉回和推出给出本体合并的范畴论解释,其中文献[16]进一步给出相应的推出算法和拉回算法。但这些工作侧重于研究本体间的合并关系,没有进一步分析本体的实例化及本体合并的一些重要性质。

到稿日期:2015-09-20 返修日期:2016-01-22 本文受广东省自然科学基金(2015A030310318),广东省医学科学技术研究基金项目(A2015065),国家自然科学基金资助项目(61103038)资助。

余珊珊(1980—),女,博士,讲师,CCF会员,主要研究方向为本体、范畴论、程序语言,E-mail:susyu@139.com;**苏锦钿**(1980—),男,博士,副教授,主要研究方向为大数据、形式语义、范畴论,E-mail:sujd@scut.edu.cn(通信作者)。

针对上述问题，本文首先给出本体、本体映射和本体实例化的范畴论解释，然后结合共极限和推出等概念给出本体间合并的定义，并分析了一些重要的合并关系。本文第 2 节介绍部分相关的理论；第 3 节给出本体、本体映射及本体实例化的范畴论解释；第 4 节利用本体映射给出本体间 V-对齐的定义，并结合共极限及推出分析了本体合并关系；最后总结全文并展望下一步的研究工作。

2 相关理论

本节介绍本文相关的部分范畴论知识，更详细的介绍可另行参考文献[1,17]。

2.1 范畴

范畴是用于描述某一类数学对象(如群、集合或完全偏序集)的本质的数学结构，可抽象地看成是由对象及对象间满足一定性质的射组成的。

定义 1[1,17] 一个范畴 $\boldsymbol{C}$ 定义为：

1)一组对象的集合，记为 $obj(\boldsymbol{C})$。

2)任意对象 $X,Y(X,Y\in obj(\boldsymbol{C}))$ 间的关系用集合 $\boldsymbol{C}(X,Y)$ 表示。$\boldsymbol{C}(X,Y)$ 中的元素称为射，且满足条件：$\boldsymbol{C}(X,Y)\cap\boldsymbol{C}(X',Y')=\emptyset$, if $X\neq X'$ 或 $Y\neq Y'$。

每一个范畴 $\boldsymbol{C}$ 同时满足以下条件：

1)组合律：对于对象 $X,Y,Z\in obj(\boldsymbol{C})$ 以及射 $f\in\boldsymbol{C}(X,Y)$ 和 $g\in\boldsymbol{C}(Y,Z)$，存在唯一的射 $g\circ f\in\boldsymbol{C}(X,Z)$，称为 f 与 g 的组合；

2)结合律：对于对象 $X,Y,Z\in obj(\boldsymbol{C})$ 以及射 $f\in\boldsymbol{C}(X,Y)$、$g\in\boldsymbol{C}(Y,Z)$ 和 $h\in\boldsymbol{C}(Z,W)$，有 $h\circ(g\circ f)=(h\circ g)\circ f$；

3)标识射：所有对象 $X\in obj(\boldsymbol{C})$ 上均存在射 $Id_X\in\boldsymbol{C}(X,X)$，使得对任意的 $f\in\boldsymbol{C}(Y,X)$ 和 $g\in\boldsymbol{C}(X,Y)$，有 $Id_Y\circ f=f$ 和 $g\circ Id_X=g$。

由于本体就是描述概念及概念间的关系，即本质上可抽象地看成是由对象及对象间的关系所构成的，因此利用范畴论作为本体描述的数学理论基础能够自然直观地刻画本体及其关系。

给定范畴中的对象和射，可定义其上的各种基本构造，典型的有积(Product)、共积(Coproduct)、回拉(Pullback)、推出(Pushout)、极限(Limit)和共极限(Colimit)等。

2.2 函子和自然转换

在范畴的基础上，可进一步定义函子，用于描述不同范畴间的关系。

定义 2[1,17] 任意两个范畴 $\boldsymbol{C}$ 和 $\boldsymbol{D}$ 之间的函子 $F:\boldsymbol{C}\to\boldsymbol{D}$ 是指满足以下条件的指派：

1)对范畴 $\boldsymbol{C}$ 中的每一个对象 X，范畴 $\boldsymbol{D}$ 中都指派一个对象 FX 与之对应；

2)对范畴 $\boldsymbol{C}$ 中的每一个射 $f:X\to Y$，范畴 $\boldsymbol{D}$ 中都指派一个射 $Ff:FX\to FY$ 与之对应。且 F 同时满足：

①对范畴 $\boldsymbol{C}$ 中的每一个对象 X，有 $FId_X=Id_{FX}$；

②对范畴 $\boldsymbol{C}$ 中任意两个可组合的射 $f:X\to Y$ 和 $g:Y\to Z$，有 $F(g\circ f)=Fg\circ Ff$。

函子给出了两个范畴中各个对象和射之间的映射关系，且同时保持原范畴中的标识射及射之间的组合关系，即函子就是保持范畴结构的指派。

对于函子，可定义自然转换，用于描述它们间的转换关系。

定义 3[1,17] 给定两个函子 $F,G:\boldsymbol{C}\to\boldsymbol{D}$，从 F 到 G 的自然转换 $\lambda:F\Rightarrow G$ 定义为：对于范畴 $\boldsymbol{C}$ 中的每一个对象 X，都有范畴 $\boldsymbol{D}$ 中的一个射 $\lambda_X:F_X\to G_X$ 与之对应，且对于范畴 $\boldsymbol{C}$ 中的任意射 $f:X\Rightarrow Y$，范畴 $\boldsymbol{D}$ 中都存在等式：$Gf\circ\lambda_X=\lambda_Y\circ Ff$，即使得图 1 满足图表交换条件。

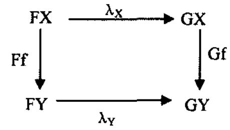


图 1 自然转换的定义

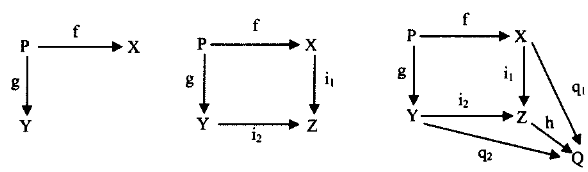
自然转换实际上就是保持范畴结构的自然的转换，其中自然的意思是指它与具体的对象类型无关。例如，在程序语言中自然转换可看成是与类型无关的类属操作，而在本体描述中自然转换可看成是与具体描述语言无关的转换操作。

范畴论中的许多概念和性质证明都可以通过图表的形式进行描述，特别是利用图表的交换性。所谓的图表交换是指一个图表中任意有相同源点和目标顶点的不同有向路径对应着范畴中相同的射，即若由相同的源点经不同的路径可到达相同的目标顶点，那么它们可以看成是对应着范畴中的相同射，例如由图 1 中图表交换可得到等式 $Gf\circ\lambda_X=\lambda_Y\circ Ff$。

2.3 推出和回拉

推出是一种典型的范畴论构造，其定义为如下。

定义 4[1,17] 设 $J=(\cdot\leftarrow\cdot\rightarrow\cdot)$ 为范畴 $\boldsymbol{C}$ 中的一个 J 型图(如图 2(a)所示)，若图 2(a)的共极限存在，则该共极限可看成是 $\boldsymbol{C}$ 中一个交换图表(如图 2(b)所示)，使得对于任意对象的 Z 和所有满足 $i_1\circ f=i_2\circ g$ 的射，均存在唯一射 $h:Z\to Q$，且满足 $q_1=h\circ i_1$ 和 $q_2=h\circ i_2$，即图 2(c)中的所有图表均满足图表交换条件。则称图 2(b)为范畴 $\boldsymbol{C}$ 的一个推出，其中 i_1 称为 g 沿 f 的推出，i_2 称为 f 沿 g 的推出。

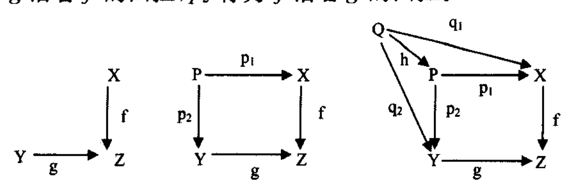


(a) (b) (c)

图 2 共极限和推出

对偶地，可给出回拉的范畴论定义。

定义 5[1,17] 设 $J=(\cdot\rightarrow\cdot\leftarrow\cdot)$ 为范畴 $\boldsymbol{C}$ 中一个 J 型图(如图 3(a)所示)，若图 3(a)中的极限存在，则该极限可看成是 $\boldsymbol{C}$ 中的一个交换图表(如图 3(b)所示)，使得对任意的对象 Q 和所有满足 $g\circ q_2=f\circ q_1$ 的射，均存在唯一射 $h:Q\to P$，且满足 $q_1=p_1\circ h$ 和 $q_2=p_2\circ h$，即图 3(c)中的所有图表均满足交换条件。则称图 3(b)为范畴 $\boldsymbol{C}$ 的一个回拉，其中 p_1 称为 g 沿着 f 的回拉，p_2 称为 f 沿着 g 的回拉。



(a) (b) (c)

图 3 极限和回拉

推出和回拉构成了范畴对偶概念，且可分别看成是范畴

论中的共极限和极限的一种特殊情况。范畴论中的初始对象、二元共积和共等值子等都是共极限的一些典型实例，而终结对象、二元积和等值子都是极限的一些典型实例。而且，共极限和极限本身也是互为范畴对偶概念。

在本体研究中，推出可用于表示本体间的最小特殊化或包含最小的复杂性，而回拉可用于表示本体间的最大抽象化或包含最多的共同部分。对范畴论中所有的图表来说，共极限总是存在的，因此推出非常适合用于描述本体间的集成。例如文献[2,14-16]的研究均采用了这种思路。不过值得注意的是，对范畴论中所有的图表来说，回拉却不一定总是存在。

3 本体范畴

3.1 本体及本体映射

目前，本体存在多种不同的形式化定义，但一般都包含概念集、关系集及相应的公理语义等部分。对范畴论来说，其侧重于描述本体及本体间的关系，而非本体的内部结构，因此本文采用文献[10]中的本体形式化定义，并暂时忽略本体上的公理语义部分。

定义 6[10] 本体 O 可描述为一个四元组(T,Rel,H,R)，其中 T 表示概念集合，其元素 $c(c\in T)$为表示特定领域中一组实体或事物的概念；Rel 表示关系集合，其元素 $r(r\in Rel)$表示概念与概念或属性与属性之间的关系；Rel 包括两部分：分类(taxonomies)关系 $H\subseteq T\times T$ 和连接关系 $R\subseteq Rel\to T\times T$，其中 H 表示概念与概念之间的父类和子类等上下层关系，R 表示除上下层次关系外的其他关系。

在本体的基础上可进一步定义本体间的映射。

定义 7[10] 本体 $O_1=(T_1,Rel_1,H_1,R_1)$和 $O_2=(T_2,Rel_2,H_2,R_2)$间的映射 $map:O_1\to O_2$ 为二元组$\langle map_1,map_2\rangle$，其中 $map_1:T_1\to T_2$，$map_2:Rel_1\to Rel_2$，且满足条件：1)$c\in T_1\Rightarrow map_1(c_1)\in T_2$；2)$(c_1,c_2)\in H_1\Rightarrow(map_1(c_1),map_1(c_2))\in H_2$；3)$(c_1,c_2)\in R_1(Rel_1)\Rightarrow(map_1(c_1),map_1(c_2))\in R_2(map_2(Rel_1))$。

本体映射给出了两个本体中的概念及其关系间的对应关系，即将源本体中的概念映射到目标本体的对应概念上，且同时保持源本体中的层次关系及其他关系。典型地，若 O_1 和 O_2 满足 $T_1\subseteq T_2$、$Rel_1\subseteq Rel_2$、$H_1\subseteq H_2$ 和 $R_1\subseteq R_2$，即 $map:O_1\to O_2$ 为入(injective)射，则称 O_1 为 O_2 的一个子本体，记为 $O_1\subseteq O_2$。

下面以两个医药本体为例来说明如何利用范畴论给出本体及本体映射的描述。

例 1 给定两个本体 $O_1=(T_1=\{$西药，麻醉类，维生素类，补钙补锌类$\}$，$Rel_1=\{H_1,R_1\}$，$H_1=\{$(麻醉类，西药)，(维生素类，西药)，(补钙补锌类，维生素类)，(补钙补锌类，西药)$\}$，$R_1=\{(r_1$，麻醉类，维生素类$)$，$(r_2$，补钙补锌类，麻醉类$)\})$和 $O_2=\{T_2=\{$中西成药，处方药，非处方药，口服类，外用类$\}$，$Rel_2=\{H_2,R_2\}$，$H_2=\{$(处方药，中西成药)，(非处方药，中西成药)，(口服类，非处方药)，(外用类，非处方药)，(口服类，中西成药)$\}$，$R_2=(s_1$，处方药，非处方药$)$，$(s_2$，口服类，处方药$)\}$，它们之间的映射 $map=\langle f,g\rangle:T_1\times Rel_1\to T_2\times Rel_2$，如图 4 所示。其中，$f$ 将 O_1 中的每一个概念 $c\in T_1$ 映射到 O_2 中的对应概念 $f(c)\in T_2$，且保持概念间的层次结构，即$(c_1,c_2)\in H_1\Rightarrow(f(c_1),f(c_2))\in H_2$。而射 g 则将 O_1 中的每一个其他关系 $r\in R_1$ 映射为 O_2 中的对应关系，即$(r,c_1,c_2)\in R_1\Rightarrow(g(r),f(c_1),f(c_2))\in R_2$，如 $g(r_1)=s_1$ 和 $g(r_2)=s_2$。

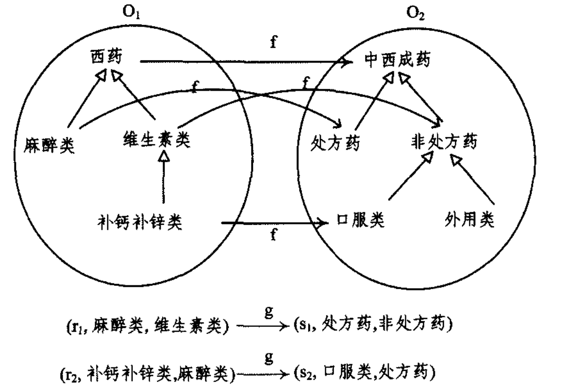


图 4 本体 O_1 和 O_2 间的映射

在许多问题域中，本体间的映射往往不一定是严格相等的，而是满足一定的语义相似度。文献[18]总结了一些常见的语义相似度度量方法，并分别针对单本体和多本体中的语义相似度计算进行了分析。

利用对象、射及范畴可从不同的抽象层次刻画本体及本体间的关系。例如，当利用某种本体语言(如 OWL)对问题域中不同本体及其关系进行建模时，可将每一个本体看成是对象，将本体间的映射看成是对象间的射，则由本体及它们间的映射可构成该语言下的一个本体范畴，且该范畴描述了同一本体语言下的不同本体及它们间保持概念关系的映射。

在本文中，将由本体及它们间的映射所构成的本体范畴记为 *Onto*，且如无特别说明均假定范畴 *Onto* 存在二元积和二元共积。

类似地，若将一个本体的所有状态集看成是不同的对象，将状态间的变迁或触发这些变迁的操作看成是射，则可构成该本体的一个状态范畴，且该范畴描述了同一个本体在不同消息下的状态变迁或演化关系。因此，利用范畴论可以从不同的抽象层次对本体进行刻画，并通过函数或自然转换描述它们之间的转换关系。

在范畴的基础上可结合范畴论中的许多抽象概念(如自然转换、伴随)及范畴构造(如积、共积、推出、拉回等)对本体及本体间的关系进行描述、推理和验证。例如，本体中的类、属性或关系更改了名称，或者增加/删除一个类、属性、泛化或关联关系。对本体来说，更改本体本身或其中的类、属性或关系的名称并不会改变其结构和语义关系，这意味着更改前、后源本体和目标本体之间存在一一对应关系，即相当于存在一个同构射。若在本体 O_1 中添加一个类、属性或关系并构成一个新的本体 O_2，则意味着本体 O_2 可看成是在 O_1 的基础上增加类、属性或关系，且同时保留原有的类、属性及关系，因此可表示为一个入射 $f:O_1\to O_2$。若在本体 O_1 中删除一个类、属性或关系并构成一个新的本体 O_2，则可表示为一个投影射。典型地，左投影射 $fst:O_1\times O_2\to O_1$ 或右投影射 $snd:O_1\times O_2\to O_2$ 可看成是将本体 $O_1\times O_2$ 中的 O_1(或 O_2)删除。

给定两个不同的本体范畴 $Onto_1$ 和 $Onto_2$，可以利用函子 $F:Onto_1\to Onto_2$ 进一步描述它们之间的关系。例如图 5 中函子 $F:Onto_1\to Onto_2$ 将范畴 $Onto_1$ 中的本体映射为 $Onto_2$ 中对应的本体，且保持本体间的映射及映射间的组合关系。

典型地，若将 $Onto_1$ 和 $Onto_2$ 分别看成同一概念模型在不同本体语言(如 OWL 和 RDFS)描述下所构成的本体范畴，

则函子相当于描述了该概念模型在不同语言间的转换关系。

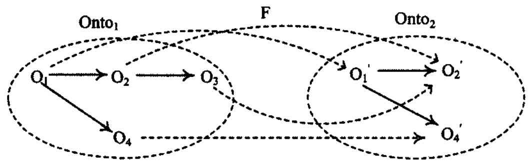


图 5　本体范畴间的函子例子

3.2　本体实例化

给定具体的问题域，本体中的各个概念在其生命周期中往往实例化成各个不同具体的对象集合。为了描述相同或不同本体的对象间的关系，下面利用函子进一步给出本体实例化的定义。

定义 8　给定本体范畴 *Onto*，称 $I:Onto\rightarrow Set$ 为 *Onto* 的实例化函子，当且仅当满足：1）对 *Onto* 中任意一个对象 $O\in Obj(Onto)$ 都存在相应的集合 $I(O)\in Obj(Set)$；2）对 *Onto* 中的任意一个射 $f\in \mathbf{C}(X,Y)$ 都存在相应的函数 $I(f)$，且保持范畴 *Onto* 中的组合律、结合律和标识射。

对范畴 *Onto* 来说，$I(O)$ 可看成是本体 O 中各个概念的实例化对象集合，并同时将各个概念间的关系映射为对应的实例化对象间的函数关系。相应地，I 将各个本体间的关系映射为对应的实例化对象集合间的函数关系。

在定义 8 的基础上，可以利用自然转换进一步描述实例化对象间的转换关系。

给定本体范畴 *Onto* 的两个实例化函子 $I,J:Onto\rightarrow Set$，若存在自然转换 $\lambda:I\Rightarrow J$，则对于 *Onto* 中任意一个本体 $O\in Obj(Onto)$，$\lambda(O):I(O)\rightarrow J(O)$ 给出了实例化对象 $I(O)$ 到 $J(O)$ 的映射关系，即可看成是同一本体 O 的两个不同实例化对象之间的转换关系。

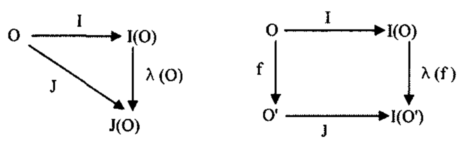


图 6　本体实例化函子间的自然转换

给定本体范畴 *Onto* 中的任意一个本体 O，则由 *Onto* 上的所有实例化函子及它们间的自然转换可构成 O 的一个实例化范畴，通常称为 O 下的共切片范畴(*Coslice Category under O*)，记为 O/Set(如图 7 所示)。则范畴 O/Set 给出了任意一个本体 O 的各个不同实例化对象及其转换关系。对于不同的本体 O_1 和 O_2 及相应的共切片范畴 O_1/Set 和 O_2/Set，若存在射 $f:O_1\rightarrow O_2$，则可定义一个射 $F_f:O_1/Set\rightarrow O_2/Set$，使得 $F_f(I_i(O_1))=I_i(f(O_1))$ 且 $F_f(\lambda_i(O_1))=\lambda_i(O_2)$，容易验证 F_f 为一个函子，且 F_f 给出了两个存在映射关系的不同本体在同一实例化函子下的不同实例化对象间的关系。

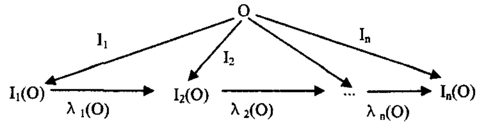


图 7　本体 O 的共切片范畴

相反地，给定本体范畴 *Onto* 中的多个本体 $O_1,O_2,\cdots,O_n$，$f_i:O_i\rightarrow O_{i+1}$ 分别为各个本体间的射，$I_i:Onto\rightarrow Set$ 分别为对应各个 O_i 上的实例化函子，其中 $1\leqslant i\leqslant n$。若对于范畴 *Set* 给定的某个实例化对象 P，均有 $I_i(O_i)=P$，满足 $I_i=I_{i+1}\circ f_i$，则由 O_i 及相应的 f_i 构成了 P 在 I_i 上的一个范畴，称为 P 上的切片范畴(Slice Category under P)，并记为 $Onto/P$(如图 8 所示)。范畴 $Onto/P$ 描述了每一个实例化对象 P 所属的各个不同抽象层次的本体及其关系。例如，典型地，$O_1,O_2,\cdots,O_n$ 之间两两构成父类和子类关系，f_i 为对应表示 O_i 和 O_{i+1} 间继承关系的射。

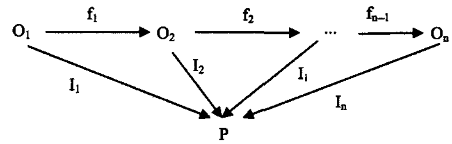


图 8　实例化对象 P 的切片范畴

值得注意的是，对于不同的实例化对象 P_1 和 P_2 及相应的切片范畴 $\mathbf{C}/P_1$ 和 $\mathbf{C}/P_2$，即使存在射 $f:P_1\rightarrow P_2$ 且满足 $I_i(O_i)=P_1$，也不一定可确定一个对应的函子 $F_f:\mathbf{C}/P_1\rightarrow \mathbf{C}/P_2$。

从上面可以看出，利用范畴论可以将本体的不同实例化对象或者实例化对象所属的不同抽象层次的本体描述成对应的共切片范畴或切片范畴，并通过函子进一步刻画它们间的转换关系。

4　本体合并

本体合并[2]是本体研究中的一个热点，主要研究如何将多个本体按照某种语义关系或约束条件合并成一个更复杂的本体。从形式化描述的角度来看，本体合并通常是指在本体映射的基础上，将 $n(n\geqslant 2)$ 个相关的本体合并成一个新的本体的过程。而新本体是这 n 个本体的并集，不仅包含原来 n 个本体的语义相似部分，也包含了语义不同的部分。

为了描述本体间的各种合并，首先需要确定本体间的映射关系。例如，若范畴 *Onto* 中的本体映射 $f:O\rightarrow O'$ 为左注入射 $inl:O_1\rightarrow O_1+O_2$ 或右注入射 $inr:O_2\rightarrow O_1+O_2$，则意味着本体 O' 是由 O_1 和 O_2 的不相交并(即二元共积)所得到的一个新本体 O_1+O_2。因此，O_1+O_2 给出了两个本体 O_1 和 O_2 间的一种合并，相当于 OWL 中的 owl:unionOf。

对于本体来说，有时不一定存在直接的映射关系，而是具有某种间接关系，例如本体中的概念或关系存在交集。为了描述这一类的合并，下面参照文献[2,18,19]给出本体对齐的定义。

定义 9　对于范畴 *Onto* 中的本体 O、O_1 和 O_2，若映射 $f:O\rightarrow O_1$ 和 $g:O\rightarrow O_2$ 存在相应的推出，则称 $\langle f,g\rangle:O\rightarrow O_1\times O_2$ 为 O 对 O_1 和 O_2 的 V-对齐，或简记为 $\langle O,f,g\rangle$。

V-对齐 $\langle O,f,g\rangle$ 利用本体间的映射 f 和 g 给出了两个本体 O_1 和 O_2 间的一种间接对应关系，其中 O 是 O_1 和 O_2 中概念、属性或关系间的语义交集部分。

对于给定的本体 O、O_1 和 O_2 及相应的 V-对齐 $\langle O,f,g\rangle$，推出 $[f',g']:O_1+O_2\rightarrow O'$ 给出了 O_1 和 O_2 的一个合并关系，记为 $[f',g',O']$，其中 $f'\circ f=g'\circ g$。即本体 O' 可看成是在 O 的基础上添加了 O_1 和 O_2 中除 O 外的其他部分。由共极限的泛性质可知，对于 O_1 和 O_2 间所有其他的合并 O''，都存在一个从 O' 到 O'' 的唯一射，因此 O' 为 O_1 和 O_2 间的一种最小合并。再由推出的存在性和唯一性可知，这种最小合并不仅存在，而且是唯一的。文献[16]给出了基于推出的本体合并算法。

相对于其他很多数学理论，利用范畴论中的共极限和推

出来描述本体间的合并关系具有更大的优势。例如，OWL 本体范畴本身就是共完备的，即该范畴中存在所有的推出[20]，这意味着 OWL 本体间的合并都可以直接利用推出和共极限进行描述，而且这种描述是合理的，不需要额外证明其存在性和唯一性。

例 2 给定本体 $O=(T=\{$西药，麻醉类$\}, Rel=\{H,R\}, H=\{($麻醉类，西药$)\}, R=\varnothing\})$，$O_1=(T_1=\{$西药，麻醉类，神经类，镇痛药$\}, Rel_1=\{H_1,R_1\}, H_1=\{($麻醉类，西药$)$，$($神经类，西药$)$，$($镇痛药，神经类$)\}$，$R_1=\{(r_1$，麻醉类，神经类$)$，$(r_2$，镇痛药，麻醉类$)\}$ 和 $O_2=\{T_2=\{$西药，麻醉类，局部麻醉类$\}, Rel_2=\{H_2,R_2\}, H_2=\{($麻醉类，西药$)$，$($局部麻醉类，麻醉类$)\}, R_2=(\{s_1$，局部麻醉类，西药$)\}$，以及相应的 V-对齐 $\langle O, f,g\rangle: O\rightarrow O_1\times O_2$。则由推出可得到相应的本体 $Z=\{\{$西药，麻醉类，局部麻醉类，神经类，镇痛药$\}, Rel'=\{H', R'\}, H'=\{($麻醉类，西药$)$，$($神经类，西药$)$，$($镇痛药，神经类$)$，$($局部麻醉类，麻醉类$)\}, R'=(\{z_1$，麻醉类，神经类$)$，$(z_2$，镇痛药，麻醉类$)$，$(z_3$，局部麻醉类，西药$)\}\}$，如图 9 所示。

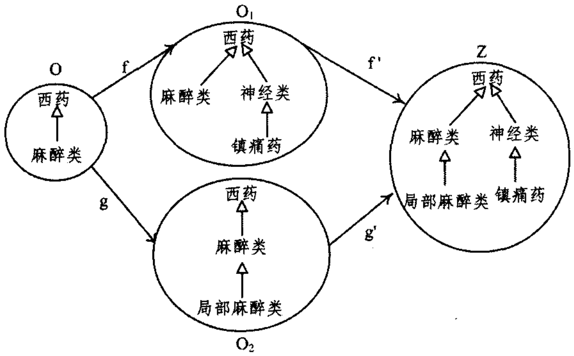


图 9 基于 V-对齐的本体合并例子

定理 1 范畴 *Onto* 中的 V-对齐满足以下组合关系：

1）给定射 $h: O\rightarrow O'$ 和 V-对齐 $\langle f,g\rangle: O'\rightarrow O_1\times O_2$，则 $\langle O, f\circ h, g\circ h\rangle$ 也为 V-对齐（如图 10(a)所示）；

2）给定 V-对齐 $\langle f,g\rangle: O\rightarrow O_1\times O_2$ 和 $\langle g,h\rangle: O\rightarrow O_2\times O_3$，则 $\langle O,f,h\rangle$ 也为 V-对齐（如图 10(b)所示）；

3）给定 V-对齐 $\langle f,g\rangle: O\rightarrow O_1\times O_2$ 和射 $h: O_1\rightarrow O_3$，则 $\langle O,h\circ f,g\rangle$ 也为 V-对齐（如图 10(c)所示）。

4）给定 V-对齐 $\langle f,g\rangle: O\rightarrow O_1\times O_2$ 和 $\langle h,i\rangle: O_1\rightarrow O_3\times O_4$，则 $\langle O,h\circ f,g\rangle$ 和 $\langle O,i\circ f,g\rangle$ 也均为 V-对齐（如图 10(d)所示）。

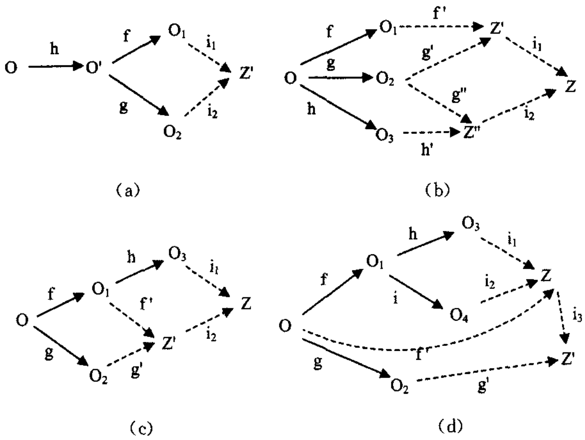


图 10 V-对齐组合

证明：1）由 V-对齐的定义可知，$\langle O', f,g\rangle$ 存在相应的推出 $[i_1,i_2, Z]$，且满足 $i_1\circ f=i_2\circ g$。由射之间的组合关系可得 $i_1\circ f\circ h=i_2\circ g\circ h$。再由推出的泛性质可知，对于所有的 R 及满足条件 $r_1\circ f=r_2\circ g$ 的射 $r_1: O_1\rightarrow R$ 和 $r_2: O_2\rightarrow R$，均存在唯一的射 $h': Z\rightarrow R$，使得 $r_1=h'\circ i_1$ 和 $r_2=h'\circ r_2$ 成立，如图 11 所示。显然，$\langle Z, f\circ h, g\circ h\rangle$ 也为推出。再由定义 9 可知 $\langle f\circ h, g\circ h\rangle$ 为 V-对齐。

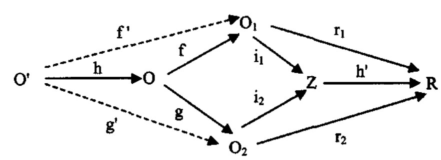


图 11 V-对齐与本体映射的组合

2）由前提条件 V-对齐 $\langle f,g\rangle$ 和 $\langle g,h\rangle$ 可知存在推出 $[f',g',Z']$ 和 $[f',g',Z'']$，且 $\langle O_2,g',g''\rangle$ 也为 V-对齐，因此推出 $[i_1,i_2,Z]$ 必然存在。再由推出的唯一性及定义 9 可知 $\langle O, f,h\rangle$ 也为 V-对齐。

3）由前提条件 V-对齐 $\langle f,g\rangle$ 可知存在推出 $[f',g',Z']$，且 $\langle O_1,h,f'\rangle: O_1\rightarrow O_3\times Z'$ 也为 V-对齐，因此可得到推出 $[i_1, i_2,Z]$。再由推出的唯一性及定义 9 可知 $\langle O,h\circ f,g\rangle$ 也为 V-对齐。

4）由前面(1)的证明可知 $\langle O,h\circ f,i\circ f\rangle$ 为 V-对齐，因此 $\langle O,i\circ f,g\rangle$ 也为 V-对齐。再由(2)的证明可知 $\langle O,h\circ f,g\rangle$ 也为 V-对齐。证毕。

定理 2 给定范畴 Onto 中的 V-对齐 $\langle f,g\rangle: O'\rightarrow O_1\times O_2$ 和 $\langle h,k\rangle: O''\rightarrow O_2\times O_3$，若射 g 和 h 存在拉回 O 及相应的射 $f'': O\rightarrow O'$ 和 $k'': O\rightarrow O''$，则 $\langle O, f\circ f'', k\circ k''\rangle$ 也为 V-对齐，如图 12 所示。

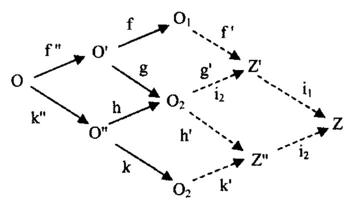


图 12 基于拉回的 V-对齐组合

证明：由前提 g 和 h 存在拉回的前提条件可得到对象 O 及相应的射 $\langle f'',k''\rangle: O\rightarrow O'\times O''$。再由前提 V-对齐 $\langle O',f,g\rangle$ 和 $\langle O'',h,k\rangle$ 可得到推出 $[f',g',Z']$ 和 $[h',k',Z'']$，且 $\langle O_2,g', h'\rangle$ 也为 V-对齐，因此可得到推出 $[i_1,i_2,Z]$。再由推出的唯一性及定义 9 可知 $\langle O, f\circ f'',k\circ k''\rangle$ 为 V-对齐。证毕。

显然，定理 1 中的 2）可看成是定理 2 的一种特殊情况，即为 O' 和 O'' 相同的情况。值得注意的是，对于范畴 *Onto* 来说，推出总是存在，因此定理 1 总是成立。但对任意的范畴论来讲，拉回却不一定存在[8]，因此定理 2 不一定总是成立。容易验证定理 2 中的 V-对齐组合满足交换律和结合律。

结束语 相对于集合论而言，范畴论具有更好的抽象性，并且正逐步取代集合论成为计算机科学中新的数学理论基础。因此，将范畴论引入到本体研究中，一方面可用于对本体进行形式化描述，从而抽象地刻画本体及本体间的各种关系，并作为本体合并、演化、推理及验证的数学理论基础；另一方面便于将基于集合论的本体研究及本体描述方法扩展到其他的范畴中，并研究不同的逻辑系统及本体表示语言间的关系及转换。

（下转第 61 页）

Conf/11-WP/6)[C]//国际民航组织第 11 次航行会议论文集. 蒙特利尔:国际民航组织,2003:7

[14] Zhou Qi, Gu Wen-zhe, Li Jing-lin, et al. A Topology Aware Routing Protocol Based ADS-B System for Aeronautical Ad Hoc Networks[C]//2012 8th International Conference Wireless Communications, Networking and Mobile Computing (WiCOM). Shanghai, China, 2012:21-23

[15] Zhang Hai, Li Gang, Chen Guang-xiao, et al. An ADS-B based clustering algorithm in aviation Ad-Hoc networks[J]. Application of Electronic Technique, 2013, 39(7):89-92(in Chinese)
张海,李纲,陈广晓,等. 基于 ADS-B 报文的航空自组网分簇算法[J]. 电子技术应用,2013,39(7):89-92

[16] Broyles D, Jabbar A, Sterbenz J P G. Design and Analysis of a 3-DGauss-markov Mobility Model for Highly-dynamic Airborne Networks[C]//International Telemetering Conference. Las Vegas, NV, 2009

[17] Huang Fei, Zhang Jun, Zhu Yan-bo, et al. Modeling and simulation of an aeronautical subnetwork based on universal access transceiver[C]//Asia Simulation Conference-7th International Conference on System Simulation and Scientific Computing ICSC2008. 2008:541-544

[18] Zheng Wei-ming. Research and Simulation of OLSR routing protocol[D]. Chengdu: University of Electronic Science and Technology of China, 2011(in Chinese)
郑伟明. OLSR 路由协议研究及仿真[D]. 成都:电子科技大学,2011

(上接第 46 页)

下一步将一方面继续研究本体间的各种合并关系及其范畴论解释,特别是推出和共极限在本体合并中的应用及其合并算法;另一方面将同时研究基于范畴论的本体演化。

参 考 文 献

[1] Awodey S. Category Theory [M]. Oxford University Press, 2006:265- 290

[2] Hitzler P, Krötzsch M, Ehrig M, et al. What is Ontology Merging? - a Category Theoretic Perspective Using Pushouts[C]//Proc. First International Workshop on Contexts and Ontologies: Theory, Practice and Applications(C&O). AAAI Press, 2005:104-107

[3] Yu Shan-shan, Li Shi-xian, Su Jin-dian. Hylomorphisms with Parameters and Its Associated Calculational Laws [J]. Journal of Computer Research and Development, 2012, 50(3):502-618 (in Chinese)
余珊珊,李师贤,苏锦钿. 一种带参数的 Hylomorphisms 及其计算律[J]. 计算机研究与发展,2012,50(3):502-618

[4] Yu Shan-shan, Li Shi-xian, Su Jin-dian. Final Coalgebraic Semantics for Behavioral Equality of Objects[J]. Computer Science, 2012, 39(2):182-190(in Chinese)
余珊珊,李师贤,苏锦钿. 对象行为等价的终结共代数语义[J]. 计算机科学,2012,39(2):182-190

[5] Yu Shan-shan, Li Shi-xian, Su Jin-dian. Formal Semantics of Object Oriented Methods Based on Coalgebras[J]. Computer Science, 2011, 38(8):142-146(in Chinese)
余珊珊,李师贤,苏锦钿. 一种基于共代数的面向对象形式语义[J]. 计算机科学,2011,38(8):142-146

[6] Su Jin-dian, Yu Shan-shan. Bialgebraic Structures of Abstract Data Types and its Calculational Laws[J]. Journal of Computer Research and Development, 2012, 49(8):1787-1803(in Chinese)
苏锦钿,余珊珊. 抽象数据类型的双代数结构及其计算定律[J]. 计算机研究与发展,2012,49(8):1787-1803

[7] Su Jin-dian, Yu Shan-shan. Corecursive operations with parameters and the associated calculational laws [J]. Journal of Computer Research and Development, 2013, 50(12):2672-2690(in Chinese)
苏锦钿,余珊珊. 带参数的共递归操作及其计算定律[J]. 计算机研究与发展,2013,50(12):2672-2690

[8] Healy M J. Category Theory as a Mathematics for Formalizing Ontologies[M]//Theory and Applications of Ontology: Computer Applications. 2010:487-510

[9] Zhang Yuan, Li Shi-xian. Research on Formal Categorical Ontologies[J]. Computer Science, 2006, 33(9):1-3(in Chinese)
章远,李师贤. 基于范畴论的形式化本体研究[J]. 计算机科学,2006,33(9):1-3

[10] Ye Dan-dan, Wang Hai-tao. Web Ontology Description Research Based on Category Theory[J]. Journal of Changchun University of Science and Technology(Natual Science Edition), 2011, 34(3):146-148,151(in Chinese)
叶丹丹,汪海涛. 基于范畴论的 Web 本体论描述研究[J]. 长春理工大学学报(自然科学版),2011,34(3):146-148,151

[11] Bench-Capon T J M, Malcolm G. Formalising Ontologies and Their Relations[C]//Proc. 10th Database and Expert Systems Applications(DEXA'99). 1999:250-259

[12] Zimmermann A, Krötzsch M, Euzenat J, et al. Formalizing Ontology Alignment and its Operations with Category Theory[C]//Proc. of the Fourth International Conference (FOIS 2006). 2006:277-288

[13] Abbas M A, Berio G. Creating Ontologies Using Ontology Mappings: Compatible and Incompatible Ontology Mappings[C]//Proc. Web Intelligence/IAT Workshops. 2013:143-146

[14] Goguen J. Three Perspectives on Information Integration[C]//Semantic Interoperability and Integration, Dagstuhl Seminar Proceedings 04391. 2005

[15] Kent R E. Semantic Integration in the Information Flow Framework[C]//Semantic Interoperability and Integration, Dagstuhl Seminar Proceedings 04391. 2005

[16] Cafezeiro I, Haeusler E H. Semantic Interoperability via Category Theory[C]//26th International Conference on Conceptual Modelling-ER Auckland. New Zealand, 2007:197-202

[17] Pierce B. Basic Category Theory for Computer Scientists[M]. The MIT Press, 1991

[18] Bench-Capon T J M, Malcolm G. Formalising Ontologies and Their Relations[C]//Proc. 10th Database and Expert Systems Applications(DEXA'99). 1999:250-259

[19] Kalfoglou Y, Schorlemmer M. Ontology Mapping: the State of the Art[J]. Knowl. Eng. Rev., 2003, 18(1):1-31

[20] Lucanu D. A Logical Foundation for the OWL Languages[C]//Proc. First International Symposium on Leveraging Applications of Formal Methods(ISoLA 2004) VolumeTR-2004-6. 2004:135-142