

程序语言中共归纳数据类型的一种 fibrations 方法

苗德成¹ 奚建清² 戴经国¹ 苏锦翎³

(韶关学院信息科学与工程学院 韶关 512005)¹ (华南理工大学软件学院 广州 510640)²

(华南理工大学计算机科学与工程学院 广州 510640)³

摘要 范畴论与共代数是程序语言中共归纳数据类型研究的传统方法,这些方法在语义行为分析与共归纳规则描述等方面存在一定的不足.针对以上问题,提出了一种 fibrations 方法以对共归纳数据类型的语义行为与共归纳规则进行研究.该方法系统分析了 fibration 上共归纳数据类型的重索引函子、对偶重索引函子与真值函子等基本逻辑结构,应用等式函子与商函子等工具建立共归纳数据类型与其语义行为在程序逻辑上的对应关系,深入分析共归纳数据类型的语义行为;并以基范畴上自函子及其在全范畴上保持等式的提升为工具构造共递归操作,抽象描述共归纳数据类型具有普适意义的共归纳规则;最后通过实例分析简要介绍了 fibrations 方法的应用.

关键词 语义行为,共归纳规则,fibrations 方法,共归纳数据类型,提升

中图法分类号 TP301.2 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.3.035

Fibrations Method of Co-inductive Data Types in Programming

MIAO De-cheng¹ XI Jian-qing² DAI Jing-guo¹ SU Jin-dian³

(School of Information Science and Engineering, Shaoguan University, Shaoguan 512005, China)¹

(School of Software, South China University of Technology, Guangzhou 510640, China)²

(School of Computer Science and Engineering, South China University of Technology, Guangzhou 510640, China)³

Abstract Traditional methods of co-inductive data types in programming, such as co-algebra and category theory, have some problems in analyzing semantics behaviors and describing co-inductive rules. Considering the status quo of co-inductive data types researches both at home and abroad, a fibrations method was proposed in this paper. Firstly, some basic logical structures of co-inductive data types are systematically analyzed over a fibration including reindexed functor, op-reindexed functor and truth functor, and the corresponding relationship between co-inductive data types and their semantic behaviors on program logic is established by equality functor and quotient functor. Then semantic behaviors of co-inductive data types are deeply analyzed. Secondly, co-recursive operations are constructed to describe abstractly co-inductive rules of co-inductive data types with universality by endo-functors in base categories and their equality-preserving lifting in total categories. At last applications of fibrations method are briefly introduced by examples.

Keywords Semantic behaviors, Co-inductive rules, Fibrations method, Co-inductive data types, Lifting

1 引言

共归纳数据类型^[1]以共代数^[2,3]为数学基础,从外部观察程序语言执行过程的动态行为.作为归纳数据类型(Inductive Data Type)的对偶概念,共归纳数据类型与归纳数据类型形成互补,共同提高程序语言语法构造与语义计算能力.目前,共归纳数据类型已成为程序语言与类型理论研究的一个重要组成部分和重点研究内容.fibrations 方法是近年来计算机科学基础理论研究的一个新兴方向,在数据库系统建模^[4,5]、软件规范^[6]和程序设计方法^[7,8]等领域有广泛的应用,特别是在范畴论研究领域,已成为前沿热点.

应用 fibrations 方法对共归纳数据类型进行研究,可以有效融合传统的范畴论与共代数研究方法,为程序语言中共归纳数据类型的语义行为分析与共归纳规则描述提供一种基于 fibrations 方法的数学框架,将共归纳数据类型融入到程序语言的形式语义与程序逻辑研究中,能提高程序语言对共归纳数据类型语义行为与共归纳规则的分析与描述能力.Hermida 与 Jacobs 在这方面做了大量的基础性研究工作^[9],特别是他们应用纤维化积与共积定义函子提升的思想为本文的研究提供了一种思路.本文应用 fibrations 方法研究程序语言中共归纳数据类型的基本思路是:以共归纳数据类型构成基范畴的对象集,以共归纳数据类型的语义行为构成全范畴的

到稿日期:2015-02-02 返修日期:2015-05-11 本文受国家自然科学基金项目(61103038),广东省自然科学基金项目(S2013010015944),广东省战略性新兴产业核心技术攻关项目(2011A010801008,2012A010701011,2012A010701003),韶关市科技计划项目(2013CX/K61)资助.
苗德成(1979-),男,博士,副教授,主要研究方向为形式语言理论、范畴论方法,E-mail:tony10860@126.com;奚建清(1962-),男,博士,教授,博士生导师,主要研究方向为数据库与网络计算;戴经国(1962-),男,硕士,教授,主要研究方向为高可靠性软件、形式语义学;苏锦翎(1980-),男,博士,副教授,主要研究方向为形式化方法、双代数.

对象集,以应用等式函子与商函子等工具建立共归纳数据类型与其语义行为在程序逻辑上的对应关系,以基范畴上自函子及其在全范畴上保持等式的提升为工具构造共归纳数据类型的共递归操作,抽象描述共归纳数据类型具有普适意义的共归纳规则。

本文的主要工作是应用 fibrations 方法研究程序语言中共归纳数据类型的语义行为及其共归纳规则。第 2 节首先引入卡式射(Cartesian Arrows)与对偶卡式射(Opcartesian Arrows)、fibration 与 opfibration 等本文研究工作所必需的基本概念;第 3 节构造关系 fibration,应用等式函子等工具分析共归纳数据类型的语义行为;在此基础上,第 4 节以自函子及其提升为工具构造共归纳数据类型上的共递归操作,抽象描述具有普适意义的共归纳规则;第 5 节分析当前共归纳数据类型研究领域的相关工作;最后总结全文并给出后续研究工作。

2 fibration 与 opfibration

2.1 fibration 与重索引函子

本文假定读者具备对象、态射与函子等范畴论基础。若范畴全体对象与态射均构成集合,则该范畴是小范畴(Small Category)^[10],本文研究对象均基于小范畴。记 $Obj C$ 为范畴 C 的对象集, $Mor C$ 为范畴 C 的态射集。下面引入卡式射、fibration 等基本概念。

定义 1^[10] 设 $P: T \rightarrow B$ 是两个小范畴 T, B 间的一个函子, $f: C \rightarrow D \in Mor B, P(Y) = D$ 。对 T 中一个态射 $u: X \rightarrow Y \in Mor T$, 如果 $P(u) = f$ 且对 $\forall v: Z \rightarrow Y \in Mor T$ 与 $\forall h: P(Z) \rightarrow C \in Mor B$, 有 $f \circ h = P(v)$, 并存在唯一的 $w: Z \rightarrow X \in Mor T$, 使 $u \circ w = v$ 且 $P(w) = h$, 则称态射 u 是 f 与 Y 的卡式射。

对 f 与 Y 的卡式射 u , 称 u 位于 f 上; 类似地, 称 Y 位于 D 上。若 u 是范畴 T 中的一个锥^[10], 则由锥态射 w 的唯一性可知, 定义 1 中的卡式射 u 是 T 中的泛锥(Universal Cone)即极限锥。相应地, 泛锥 u 的顶点 X 为 u 的终对象^[11], 而由泛锥的泛性质知卡式射 u 是一个同构(Isomorphism)。为简化问题陈述而又不失一般性, 记定义 1 中 f 与 Y 的卡式射 u 为 f_Y^\downarrow 。

定义 2^[10] 设 $P: T \rightarrow B$ 是一个函子, 如果对 $\forall Y \in Obj T$ 与 $\forall f: C \rightarrow P(Y) \in Mor B$, 都存在一个 f 与 Y 的卡式射 f_Y^\downarrow , 则称 P 是一个 fibration。

由定义 2 知, fibration 本质上是一种确保大量卡式射存在的函子, 对于一个 fibration $P: T \rightarrow B$, 称 B 为基范畴, T 为全范畴。对基范畴 B 中的一个对象 $C, \exists X \in Obj T, k \in Mor T$, 若有 $P(X) = C$ 与 $P(k) = id_C$, 则 X 与 k 构成的子范畴 T_C 称为对象 C 上的纤维(Fiber)^[10], 并称 k 为垂直态射。实际上, 纤维 T_C 是全范畴 T 的一个全子范畴(Full Subcategory)。

例 1 设 Set 为集合范畴, $\forall X \in Obj Set, X$ 上的一个谓词是一个二元组 $\langle X, P \rangle$, 其中, $P: X \rightarrow Set$ 。对 $\forall x \in X, P(x)$ 构成一个集合, 描述 x 的语义行为, 并称 X 为谓词 $\langle X, P \rangle$ 的定义域。谓词 $\langle X, P \rangle$ 到 $\langle X', P' \rangle$ 的态射是一个序对 $(f, f^-): \langle X, P \rangle \rightarrow \langle X', P' \rangle$, 其中 $f: X \rightarrow X'$ 是相应谓词定义域上的函数, 而对 $\forall x \in X, f^-: P(x) \rightarrow P'(f(x))$, 将 $P(x)$ 映射为 $P'(f(x))$ 。谓词及其态射构成谓词范畴 P , 进而得到谓词 fi-

bration $P: P \rightarrow Set$ 将全范畴 P 的对象 $\langle X, P \rangle$ 映射为 X 。

取例 1 的谓词 fibration P 基范畴 Set 中态射 $g: X \rightarrow Y$, 对 $\langle Y, Q \rangle \in Obj P$, 记 Id 为单位函子, g 与 $\langle Y, Q \rangle$ 关于谓词 fibration P 的一个卡式射为 $g_{\langle Y, P \rangle}^\downarrow = (g, Id_{sa}); Qg \rightarrow Q$ 。

例 2 文献[12]的定义域 fibration 与共域 fibration。记范畴 B 的射范畴(Arrow Category)^[10] 为 B^\rightarrow , 定义域函子 $dom: B^\rightarrow \rightarrow B$ 将 B^\rightarrow 的一个对象 $f: X \rightarrow Y$ 映射为 B 的对象 X , 称 dom 为 B 上的定义域 fibration。函子 $cod: B^\rightarrow \rightarrow B$ 将 B^\rightarrow 的一个对象 $f: X \rightarrow Y$ 映射为 B 的对象 Y , 若 B 有拉回(Pullbacks)^[10], 则称 cod 为 B 上的共域 fibration。

例 2 的共域 fibration 要求基范畴 B 有拉回, 而定义域 fibration 没有此特定要求。同时, 对 Y 上纤维 B_Y^\rightarrow 中对象 $f: X \rightarrow Y$, 有 $f': X' \rightarrow Y \in Mor B$, 则 f' 与 f 关于共域 fibration cod 的一个卡式射为 f 沿 f' 的拉回方形(Pullback Square)^[10]。

例 3 文献[9]的子对象 fibration。设范畴 B 有拉回, $Sub(B)$ 为 B 的子对象构成的范畴, 即 $Sub(B)$ 的对象为单射等价类(Equivalence Classes of Monos), 如 $[f]: X \twoheadrightarrow I \in Obj Sub(B)$, 对 $Sub(B)$ 中的另一对象 $[g]: Y \twoheadrightarrow J$, 有态射 $(I \rightarrow J): [f] \rightarrow [g] \in Mor Sub(B)$, 记为 $\alpha: I \rightarrow J, \beta: X \rightarrow Y$, 满足图表交换 $\alpha \circ [f] = [g] \circ \beta$ 。子对象 fibration S 为 $S: Sub(B) \rightarrow B$, 将一个单射等价类 $[f]$ 映射为其共域。

记 $f^*(Y)$ 为卡式射 f_Y^\downarrow 的定义域, 则 $f^*(Y)$ 位于 C 上, 即 $Y \in Obj T_D, f^*(Y) \in Obj T_C$, 由此得到重索引函子的定义。

定义 3 基范畴 B 中的态射 $f: C \rightarrow D$ 扩展为纤维间的一个函子 $f^*: T_D \rightarrow T_C$, 称 f^* 为由 f 归纳的重索引函子。

f 是基范畴中共归纳数据类型间的关联, 而重索引函子 f^* 则是 f 在全范畴上的提升, 对应着语义行为间的关联。下面引入 fibration 的对偶概念 opfibration。

2.2 opfibration 与对偶重索引函子

定义 4^[10] 设 $P: T \rightarrow B$ 是两个小范畴 T, B 间的一个函子, $f: C \rightarrow D \in Mor B, u: X \rightarrow Y \in Mor T$ 。如果 $P(X) = C, P(u) = f$ 且对 $\forall v: X \rightarrow Z \in Mor T$ 与 $\forall h: D \rightarrow P(Z) \in Mor B$, 有 $h \circ f = P(v)$, 并存在唯一的 $w: Y \rightarrow Z \in Mor T$, 使得 $w \circ u = v$ 且 $P(w) = h$, 则称态射 u 是 f 与 X 的对偶卡式射。

与定义 1 类似, 若 u 是范畴 T 中的一个共锥^[10], 则由共锥态射 w 的唯一性可知, 定义 4 中的对偶卡式射 u 是 T 中的泛共锥(Universal Cocone)即共极限共锥。相应地, 泛共锥 u 的顶点 Y 为 u 的初始对象^[11], 而由泛共锥的泛性质知对偶卡式射 u 是一个同构。

定义 5^[10] 设 $P: T \rightarrow B$ 是一个函子, 如果对 $\forall X \in Obj T$ 与 $\forall f: P(X) \rightarrow D \in Mor B$, 都存在一个 f 与 X 的对偶卡式射, 则称 P 是一个 opfibration。

定义 6^[10] 若函子 $P: T \rightarrow B$ 既是一个 fibration, 又是一个 opfibration, 则称 P 为 bifibration。

记定义 4 中 f 与 X 的对偶卡式射 u 为 f_X^\uparrow , 令 $*f(X)$ 为对偶卡式射 f_X^\uparrow 的共域, 则 $*f(X)$ 位于 D 上, 即 $X \in Obj T_C, *f(X) \in Obj T_D$ 。

定义 7 基范畴 B 中的态射 $f: C \rightarrow D$ 扩展为纤维间的一个函子 $*f: T_C \rightarrow T_D$, 称 $*f$ 为由 f 归纳的对偶重索引函子。

定理 1 设 $P: T \rightarrow B$ 是一个 bifibration, 基范畴 B 有拉

回。若对 B 中任意一个拉回方形, 自然变换 $*s \circ t^* \rightarrow g^* \circ *f$ 是一个同构, 则称 P 满足 Beck-Chevalley 条件。

证明: 设 η 为伴随函子 $*f \dashv f^*$ 的单位, ϵ_s 为伴随函子 $*s \dashv s^*$ 的共单位, 则有 $\eta = Id_{T_B}, \epsilon_s = Id_{T_C} \circ (*s \circ t^*) \circ \eta = (*s \circ t^*) \circ (f^* \circ *f)$, 而图 1 的拉回方形满足图表交换 $f \circ t = g \circ s$, 且 s 是 t^* 沿 g 的拉回。 t 是 g 沿 f 的拉回, 由重索引函子的拉回性质有 $t^* \circ f^* \cong s^* \circ g^*$, 故有 $(*s \circ t^*) \circ (f^* \circ *f) = *s \circ (t^* \circ f^*) \circ *f \cong *s \circ (s^* \circ g^*) \circ *f$, 而 $*s \circ (s^* \circ g^*) \circ *f = (*s \circ s^*) \circ (g^* \circ *f) = \epsilon_s \circ (g^* \circ *f) = (g^* \circ *f)$, 即 $*s \circ t^* \cong g^* \circ *f$, 故自然变换 $*s \circ t^* \rightarrow g^* \circ *f$ 是一个同构。证毕。

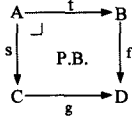


图 1 基范畴 B 中任意一个拉回方形

定理 1 实际上是由 bifibration P 的基范畴中一个拉回方形定义了全范畴 T 中各相关纤维间函子保结构的一种自然变换, 进而确保了重索引与对偶重索引函子满足合适的图表交换性质。例 1 的谓词 fibration 与例 2 的共域 fibration 均满足定理 1 的 Beck-Chevalley 条件。

3 共归纳数据类型的语义行为

3.1 关系 fibration 与等式函子

定义 8 设 $P: T \rightarrow B$ 是一个 fibration, 取 $\forall D \in Obj B$, 若 $\exists 1_D \in Obj T_D$ 为纤维 T_D 上的终对象, 且 $\forall f: C \rightarrow D \in Mor B, f^*(1_D)$ 为纤维 T_C 上的终对象, 即重索引函子 f^* 保持终对象, 则称 fibration P 有纤维化终对象。

例 1 中谓词 fibration P 的纤维化终对象是将集合 X 中所有元素均映射为单点集的函数, 例 2 中共域 fibration cod 的纤维化终对象是单位函数, 例 3 中子对象 fibration S 的纤维化终对象是单位函数的等价类。

定义 9 设 $P: T \rightarrow B$ 是一个 fibration, 函子 $T_P: B \rightarrow T$ 将 $\forall C \in Obj B$ 映射为纤维 T_C 上的终对象, 称 T_P 为 fibration P 的真值函子。

记 1_B 与 1_T 分别为基范畴 B 与全范畴 T 的终对象, 则有 $P(1_T) = 1_B$, 对 $\forall C \in Obj B$, 存在唯一的态射 $u: C \rightarrow 1_B$, 有 $T_P(C) \cong u^*(1_T)$ 。对 $\forall f: C \rightarrow D \in Mor B, f^*(T_P(D)) \cong T_P(C)$, 真值函子 T_P 将 f 映射为全范畴 T 上的卡式射 $f^*_{T_P(D)}$ 。局部单且局部满 (Full and Faithful)^[11] 的真值函子 T_P 是 fibration P 的纤维化右伴随^[7]。

定义 10^[9] 设 $P: T \rightarrow B$ 为一个 fibration, 基范畴 B 有积。令 $\Delta: B \rightarrow B$ 为 B 上的对角自函子, 将 $\forall C \in Obj B$ 映射为积对象 $C \times C$ 。 P 沿 Δ 的拉回构成 fibration $Rel(P): Rel(T) \rightarrow B$, 称 $Rel(P)$ 为 P 的关系 fibration。

$Rel(P)$ 全范畴 $Rel(T)$ 的对象为关系 (C, D) , 对另一对象 (C', D') , 令 $\alpha: C \rightarrow C', \beta: D \rightarrow D', (\alpha, \beta): (C, D) \rightarrow (C', D') \in Mor Rel(T)$ 。图 2 中的关系 fibration $Rel(P)$ 将关系 (C, D) 映射为基范畴 B 中的对象 C , 函子 Π 将 (C, D) 映射为 T 中的对象 D , 且有 $P(D) = \Delta(C)$ 。同时, 定义 10 的拉回保持性质确保 C 上关于 $Rel(P)$ 的纤维 $Rel(T)_C$ 与 $C \times C$ 上关于 P 的纤维 $T_{C \times C}$ 是同构的, 即 $Rel(T)_C \cong T_{C \times C}$ 。

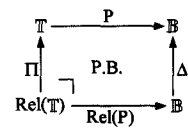


图 2 P 的关系 fibration $Rel(P)$

由给定的 fibration 通过拉回构造一个新的 fibration 的过程称为基变换 (Change of Base)^[9], 如定义 10 中 P 通过基变换构造 $Rel(P)$ 。基变换具有保持结构性质, 如保持纤维化终对象^[13], 即若 P 有真值函子 T_P , 则 $Rel(P)$ 有真值函子 $T_{Rel(P)}$, 且 $T_{Rel(P)}(C) = T_P(C \times C)$ 。如例 1 的谓词 fibration P 通过基变换构造关系 fibration $Rel(P)$, 其真值函子将集 X 映射为一个二元关系 $R: X \times X \rightarrow Set$, 即将每一个序对 (x, x') 映射为一个单点集 $\{*\}$ 。

定义 11 设 $P: T \rightarrow B$ 是一个满足 Beck-Chevalley 条件的 bifibration, B 有积, 且 T_P 为 P 的真值函子。对 $\forall C \in Obj B$, 自然变换 $\delta: Id_B \rightarrow \Delta$ 在 C 的作用函数 $\delta_C: C \rightarrow C \times C$ 扩展为对偶重索引函子 $*$ δ , 称 $Eq: B \rightarrow Rel(T)$ 为 P 的等式函子, 并记 $Eq = * \delta \circ T_P$ 。

真值函子 T_P 将 C 映射为 T_C 上的终对象 $T_P(C)$, 由定义 10 知 $Rel(P)$ 是 P 沿 Δ 的基变换, 则若 P 有纤维化终对象, 则 $Rel(P)$ 也有纤维化终对象。定义 11 中的 $* \delta$ 将 $T_P(C)$ 映射为 $* \delta(T_P(C))$, 且 $* \delta(T_P(C)) \in Obj(T_{C \times C} \cong Rel(T)_C)$ 。等式函子 Eq 将 $\forall f \in Mor B$ 映射为由 δ_f 与 $(\delta_C)^*_{T_P(C)}$ 确定的 $f \times f$ 上的唯一态射, 其直观意义是相同的参数得到相同的结果^[9]。如例 1 的谓词 fibration P , 其纤维 $Rel(T)_C$ 中的对象为等式关系 $R: X \times X \rightarrow Set$, 等式函子 Eq 将 C 映射为 $Eq(C)$ $(x, x') = 1$, 若 $x = x'$; 否则, $Eq(X)(x, x') = 0$ 。

3.2 保持等式的提升与商函子

定义 12 设 $P: T \rightarrow B$ 是一个满足 Beck-Chevalley 条件的 bifibration, B 有积, 且 P 有真值函子。令 $F: B \rightarrow B$ 是基范畴 B 上的一个自函子, 若 $Eq \circ F \cong F \circ Eq$, 则称 $F: Rel(T) \rightarrow Rel(T)$ 为 F 关于 $Rel(P)$ 的一个保持等式的提升。

对 $\forall C \in Obj B$, 在自函子 F 作用下构成一个 F -共代数 $(C, i: C \rightarrow F(C))$, C 为 i 的载体 (Carrier)。 i 与另一 F -共代数 $(D, j: D \rightarrow F(D))$ 的态射是 i 与 j 载体间的态射 $f: C \rightarrow D$, 且满足图表交换 $j \circ f = F(f) \circ i$ 。 F -共代数及其态射构成 F -共代数范畴, 记为 $Coalg_F$ 。终结 F -共代数 $(\nu F, out: \nu F \rightarrow F(\nu F))$ 若存在, 则是唯一同构的, 终结共代数的终结性泛性质所确定的唯一同构性是研究共归纳数据类型语义行为及其共归纳规则的主要工具。作为终结 F -共代数载体的共归纳数据类型, νF 是函子 F 的最大不动点 (Maximal Fixed Point), 函子 F 指共归纳数据类型 νF 的语法析构 (Destructor), out 为从外部观察 νF 在该语法析构过程中的一种语义行为。

应用等式函子 Eq 将 F -共代数 (C, i) 映射为一个 F^\perp -共代数 $(Eq(C), Eq(i): Eq(C) \rightarrow Eq(F(C)) \cong F^\perp(Eq(C)))$ 。相应地, $Eq(\nu F)$ 为终结 F^\perp -共代数的载体, 即等式函子 Eq 保持终对象。记 $Coalg(Eq)$ 为 F -共代数范畴 $Coalg_F$ 到 F^\perp -共代数范畴 $Coalg_{F^\perp}$ 的函子, 利用等式函子 Eq 将关系 fibration $Rel(P)$ 基范畴 B 中的对象与态射映射为全范畴 $Rel(T)$ 中相应的对象与态射, 通过函子 $Coalg(Eq)$ 进一步建立 F -共代数范畴 $Coalg_F$ 到 F^\perp -共代数范畴 $Coalg_{F^\perp}$ 的联系。 $out^\perp: Eq(\nu F) \rightarrow F^\perp(Eq(\nu F))$ 是关系 fibration $Rel(P)$ 全范畴 $Rel(T)$

中的一个终结 F^\perp -共代数, 则 out^\perp 是 out 在函子 $Coalg(Eq)$ 作用下的同态 (Homomorphism) 像, 即 $Coalg(Eq)(out) = out^\perp$ 。终结 F^\perp -共代数的终结性确保 out^\perp 是唯一同构的, 这种唯一同构泛性质的存在为分析共归纳数据类型的语义行为及描述其共归纳规则提供了极大的便利。

定义 13 若 fibration $P: T \rightarrow B$ 的等式函子 Eq 有一个左伴随函子 Q , 即 $Q \dashv Eq$, 且有图表交换 $F \circ Q = Q \circ F^\perp$, 则称 Q 为 P 的商函子。

基于集合论的角度, 例 1 谓词 fibration P 的商函子将关系范畴 $Rel(T)$ 中每一个关系对象映射为由该关系对象生成的最小等价类所确定的商集。若将例 3 的子对象 fibration S 的基范畴 B 限定为一个正则范畴 (Regular Category)^[10], 且 B 有共等值子 (Coequalizers)^[10], 则 S 有商函子 Q , Q 将 $Rel(S)$ 全范畴 $Rel(T)$ 中的一个等价类映射为其共等值子的共域。

与 $Coalg(Eq)$ 类似, 记 $Coalg(Q)$ 为 F^\perp -共代数范畴 $Coalg_{F^\perp}$ 到 F -共代数范畴 $Coalg_F$ 的函子, 由定义 13 伴随函子 Eq 与 Q 的伴随性质有 $Coalg(Q) \dashv Coalg(Eq)$ ^[9], 对任一 F^\perp -共代数 $(X, j; X \rightarrow F^\perp(X))$, 有 $Coalg(Q)(j) = Q(X) \rightarrow F(Q(X))$, 即 $Coalg(Q)(j) = Q(j)$, 则 $Q(j)$ 是 j 在函子 $Coalg(Q)$ 作用下的同态像, 如图 3 所示。若 $g: X \rightarrow Eq(C)$ 是 j 到 $Eq(C)$ 的 F^\perp -共代数态射, 则 $Q(j)$ 到 i 的 F -共代数态射 $h: Q(X) \rightarrow C$ 是 g 上的 F -共代数同态。类似地, g 是 h 上的 F^\perp -共代数同态。函子 $Coalg(Eq)$ 的左伴随 $Coalg(Q)$ 建立以 $Q(X)$ 为载体的 F -共代数与以 X 为载体的 F^\perp -共代数间直观的互推导关系, 为共归纳数据类型共归纳规则的形式化描述提供了一种以 νF 为终结共代数载体的简洁与一致的建模方法, 即若函子 $Coalg(Eq)$ 保持终对象, 则 F 关于 $Rel(P)$ 保持等式的提升 F^\perp 生成一个可靠的共归纳规则。

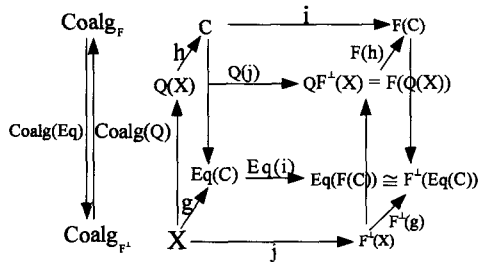


图 3 $Coalg(Q) \dashv Coalg(Eq)$ 的伴随性质

4 共归纳数据类型的共归纳规则

若共归纳数据类型定义并运用了具有等式函子与商函子的 fibration, 则其共归纳规则的形式化描述与语义行为分析是一致的。Ghani 等学者证明了以下定理成立, 该定理为文献^[12]中定理 3. 10。

定理 2^[12] 设 $P: T \rightarrow B$ 是一个满足 Beck-chevalley 条件的 bifibration, B 有积, 且 P 有真值函子与商函子。 $Rel(P): Rel(T) \rightarrow B$ 为 P 的关系 fibration, $F: B \rightarrow B$ 是基范畴 B 上的一个自函子, νF 为终结 F -共代数的载体。 F 关于 $Rel(P)$ 的每一个保持等式的提升 $F^\perp: Rel(T) \rightarrow Rel(T)$ 都有一个可靠的基于 νF 的共归纳规则。

定理 2 实际上为 F^\perp 应用终结 F -共代数在共归纳数据类型上生成共归纳规则的有效性 (Validity) 判定提供了一种可靠依据, 即若 bifibration P 定义并运用等式函子与商函子分

析共归纳数据类型的语义行为, 则其基于终结 F -共代数的共归纳规则在程序语言语义行为分析过程中是有效的。

下面在 fibrations 方法的框架内分析与描述共归纳数据类型具有普适意义的共归纳规则。首先考虑共归纳数据类型的共递归计算。基于范畴论的观点, 共归纳数据类型的共递归计算源于终结共代数语义^[2]。以共归纳数据类型为终结 F -共代数的载体 νF , 应用基范畴 B 上的自函子 F 构造共归纳数据类型的共递归操作 $unfold: (C \rightarrow F(C)) \rightarrow C \rightarrow \nu F$, 对任意一个 F -共代数 $(C, i; C \rightarrow F(C))$, 在共递归操作 $unfold$ 的作用下, $unfold i$ 将 i 映射为 i 到终结 F -共代数 out 的唯一 F -共代数态射 $unfold i: C \rightarrow \nu F$, 如图 4 所示。源于终结共代数语义的 $unfold$ 本质上是共归纳数据类型一个参数化 (Parameterized) 的共递归操作, 其共递归计算具有语义正确、扩展灵活与表达简洁等良好性质。

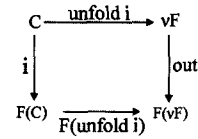


图 4 F -共代数态射

$Eq \circ F(C) \cong F^\perp \circ Eq(C)$, $Eq \circ F(\nu F) \cong F^\perp \circ Eq(\nu F)$, 而由等式函子 Eq 保持终对象性质知, $Eq(\nu F)$ 为终结 F^\perp -共代数的载体, 记 $\nu F^\perp = Eq(\nu F)$, $X = Eq(C) \in Obj Rel(T)$ 。类似地, 以 F 保持等式的提升 F^\perp 为工具构造全范畴 $Rel(T)$ 上描述共归纳数据类型语义行为的共递归操作 $unfold: (X \rightarrow F^\perp(X)) \rightarrow X \rightarrow \nu F^\perp$, 如图 5 所示, 进而对 $\forall C \in Obj B, X \in Obj Rel(T)_C$, 在 fibrations 方法的形式化框架内得到共归纳数据类型具有普适意义的共归纳规则:

$$Coind_{\nu m}: (X \rightarrow F^\perp(X)) \rightarrow X \rightarrow Eq(\nu F)$$

若 $j: X \rightarrow F^\perp(X)$ 是 F -共代数 $(C, i; C \rightarrow F(C))$ 上的 F^\perp -共代数, 则 $Coind_{\nu m} X j: X \rightarrow Eq(\nu F)$ 是 $unfold i$ 上的 F^\perp -共代数同态。

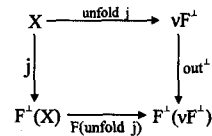


图 5 F^\perp -共代数态射

例 4 令确定有穷状态自动机 DFA 的状态空间 K 为基范畴 B 上终结 F -共代数 out 的载体 νF , Σ 为有限输入字母表, $F: K \times \Sigma \rightarrow K$ 为状态转移函数。记 1 为终对象, ϵ 为空字。对 $\forall x \in K, a \in \Sigma$, 若 $F(x, a) = 1$, 则 DFA 停机; $F(x, a) \in K$, 则 DFA 成功运行并产生一个新状态。对 DFA 中另一状态 x' , 取关系 fibration $Rel(P)$ 全范畴 $Rel(T)$ 中 DFA 任一性质 $R \in Obj Rel(T)$, 如互模拟 (Bisimulation), 则有对 R 的一个共归纳:

$$R(x, x') \text{ iff } ((F(x, \epsilon) = F(x', \epsilon)) \wedge F(x, a) \in K \wedge F(x', a) \in K \wedge R(F(x, a), F(x', a)))$$

成立。对任一 F -共代数 $(C, j; C \rightarrow F(C))$, 通过 F 关于 $Rel(P)$ 的一个保持等式提升 F^\perp , 有 F^\perp -共代数 $(X, g; X \rightarrow F^\perp(X))$, $X = Eq(C)$, 有图表交换 $(F \circ Rel(P))(X) = (Rel(P) \circ F^\perp)(X)$ 。终结 F -共代数的终结性定义 DFA 状态空间 K 上的一个共递归操作 $unfold j$, 执行 DFA 状态转换的判定; 而由终结

F^{\perp} -共代数的终结性对应得到一个共递归操作,描述 DFA 的语义行为。若 g 位于 j 上,则 $Coind_{U_{mi}} X g$ 是 $unfold j$ 上的 F^{\perp} -共代数同态,且遍历全范畴 $Rel(T)_C$ 中每一性质 $R, R \in Obj Rel(T)_C, \forall C \in Obj B$, 得到描述 DFA 行为的语义集 $\{R(X, X) | X = Eq(C)\}$ 。

互模拟是自动机理论研究的核心内容。例 4 从 fibrations 方法的角度进一步拓展了传统自动机理论的研究内容,例如在关系 fibration $Rel(P)$ 语义模型基础上建立描述 DFA 共递归计算的共归纳规则 $Coind_{U_{mi}}$, 为共归纳数据类型 DFA 的语义行为与程序逻辑提供了一种简洁的描述方式,特别是在函数式程序语言(如 Haskell)中, $Coind_{U_{mi}}$ 生成的代码片段具有易读、易写与易理解等良好性质。

5 相关研究

共归纳数据类型以共代数为数学基础,将终结性与互模拟等工具引入类型理论研究中,在程序语言动态语义行为分析与描述方面具有独特的优势。作为归纳数据类型的对偶概念,共归纳数据类型与归纳数据类型形成互补,为研究数据类型的语法构造与语义计算提供了一种可行性^[14]。从文献检索的情况看, Hagino 最早应用 Dialgebras 结构系统地研究了归纳与共归纳数据类型间的关系^[15], 为共归纳数据类型的研究奠定了基础,但在多态类型系统、程序语言的共归纳原理应用等方面仍存在着一定的不足。

众多学者的共同努力进一步推动了共归纳数据类型理论的发展。如 Nogueira 应用双代数(Bialgebra)方法研究了归纳与共归纳数据类型的关系及其在多态编程中的应用^[16]; 文献^[14]进一步利用 λ -双代数与分配律将归纳与共归纳数据类型有机融合起来,探讨了数据类型的语法构造与动态行为关系; Poll 基于子类型和继承对 Hagino 的工作进行了拓展,应用代数与共代数的对偶性质研究了归纳与共归纳数据类型的子类型和继承关系^[17]; Greiner 等学者将共归纳原理引入到程序语言的研究中,对程序语言中的共归纳数据类型进行了深入研究^[1, 18]。以上研究成果在一定程度上解决了上述问题。同时,在共归纳数据类型的应用方面, Gimenez 对共归纳数据类型在形式化理论证明工具 Coq 中的应用进行了研究^[19], Vene 对函数式程序语言 Haskell 中的共归纳数据类型进行了研究^[20]。

现有基于 fibrations 方法的研究则侧重于共归纳数据类型程序逻辑推理与共归纳规则有效性验证等方面。如 Hermida 与 Jacobs 证明了有商类型的终结共代数的共归纳规则是可靠的^[9]; Ghani 等学者^[12]在文献^[9]的基础上,突破多项式函子的局限,将其研究工作扩展为一般意义上的函子类型,但在共归纳数据类型的语义计算和程序逻辑等方面仍存在许多尚未解决的问题,如语义行为与共归纳规则的分析与描述等,特别是在共归纳数据类型的研究中共归纳规则多以自动生成为主,缺乏坚实的数学基础和精确的形式化描述。

共归纳数据类型传统的共代数研究方法在局部卡式闭范畴(Local Cartesian Closed Category)^[9]内建立类型论模型,使得共归纳数据类型与描述其语义行为的关系范畴共存于同一范畴内,导致函子与其提升是等同的,在语义行为分析与共归纳规则描述方面存在着一定的局限性。而本文在 fibrations 方法的形式化框架内展开对程序语言中共归纳数据类型的研

究,描述共归纳数据类型语义行为的关系不再局限于函数或态射,而是提升为全范畴中的对象。同时,更为重要的是,共归纳数据类型与描述其语义行为的关系范畴不再共存于同一范畴内,而是在全范畴上构造函数子提升,深入分析与抽象描述共归纳数据类型的语义行为与共归纳规则。

将 fibrations 方法应用于程序语言共归纳数据类型中的研究是传统共归纳数据类型研究方法在范畴论层面上的拓展与深化,特别是共代数方法出现后,卡式射与对偶卡式射、fibration 与 opfibration 等对偶范畴概念的有机结合,使得 fibrations 方法在程序语言共归纳数据类型中的研究呈现出强大的生命力,在计算机科学的理论研究和工程实践中具有广阔的应用前景。另外,应用 fibrations 方法对程序语言中共归纳数据类型的研究不是纯粹数学意义上的研究,而是从程序语言的应用角度出发,结合 fibrations 方法在面向对象语言、共代数规范及语义计算中的最新研究成果,对程序语言共归纳数据类型中各种核心概念的范畴性质、语义行为与规范描述等核心问题进行的基础研究。

对于融入传统程序语言设计思想的 fibrations 方法,其高度抽象性、灵活扩展性及简洁描述性的独特思路和研究方法为程序语言及其形式语义的研究带来积极和深远的影响,并极大地推动范畴论方法在计算机科学中的应用。但从目前文献检索的情况来看,国际上从事 fibrations 方法研究的学者不多,而将 fibrations 方法应用于计算机科学领域中的研究文献相对较少,尤其是针对程序语言及其形式语义展开系统、深入研究的文献更少,当前国内尚未发现其他学者对这一方法在计算机科学中的应用表示关注并进行系统的研究,希望本文的研究能够引起国内其他学者对 fibrations 方法的关注和研究。

结束语 索引共归纳数据类型(Indexed Coinductive Data Types)是一种语义计算能力更强的共归纳数据类型,可处理程序语言中更为复杂的数据结构,但当前对索引共归纳数据类型的共归纳规则的研究却很少。下一步将扩展我们的研究成果至基于离散索引对象的索引共归纳数据类型,在切片范畴(Slice Category)上应用 fibrations 方法研究索引共归纳数据类型的语义行为与共归纳规则;同时,初步探讨共归纳数据类型及其共归纳规则构成复杂形式系统的可靠性、完备性与一致性等元性质,并应用范畴论的对偶原理及适当的分配律深入分析归纳与共归纳数据类型在 fibrations 方法框架内的融合与计算。

参考文献

- [1] Greiner J. Programming with inductive and co-inductive types; Tech. Report CMU-CS-92-109[R]. Pittsburgh, USA; School of Computer Science, Carnegie Mellon University, 1992
- [2] Rutten J. Universal coalgebra: a theory of systems[J]. Theoretical Computer Science, 2000, 249(1): 3-80
- [3] Zhou Xiao-cong, Shu Zhong-mei. A survey on the coalgebraic methods in computer science[J]. Journal of Software, 2003, 14(10): 1661-1671(in Chinese)
周晓聪,舒忠梅. 计算机科学中的共代数方法的研究综述[J]. 软件学报, 2003, 14(10): 1661-1671

(下转第 212 页)

- [3] Rothermel G, Untch R H, Chu C. Test case Prioritization: An empirical study[C]// IEEE International Conference on Software Maintenance, 1999(ICSM'99). IEEE, 1999: 179-188
- [4] Cao He-ling, Jiang Shu-juan, Ju Xiao-lin. Survey of Software Fault Localization[J]. Computers Science, 2014, 41(2): 1-6 (in Chinese)
曹鹤玲, 姜淑娟, 鞠小林. 软件错误定位研究综述[J]. 计算机科学, 2014, 41(2): 1-6
- [5] Wong W E, Debroy V. A survey of software fault localization: UTDCS-45-09[R]. Department of Computer Science, University of Texas at Dallas, 2009
- [6] Jiang B, Zhang Z, Tse T H. How well do test case prioritization techniques support statistical fault localization[C]// 33rd Annual IEEE International Computer Software and Applications Conference, 2009(COMPSAC'09). IEEE, 2009, 1: 99-106
- [7] Jiang B, Chan W K. On the integration of test adequacy, test case prioritization, and statistical fault localization[C]// 2010 10th International Conference on Quality Software (QSIC). IEEE, 2010: 377-384
- [8] Gonzalez-Sanchez A, Piel é, Abreu R. prioritizing tests for software fault diagnosis [J]. Software: Practice and Experience, 2011, 41(10): 1105-1129
- [9] Gonzalez-Sanchez A, Abreu R, Gross H G. Prioritizing tests for fault localization through ambiguity group reduction[C]// 2011 26th IEEE/ACM International Conference on Automated Software Engineering (ASE). IEEE, 2011: 83-92
- [10] Xie X, Chen T Y, Kuo F C. A theoretical analysis of the risk evaluation formulas for spectrum-based fault localization[J]. ACM Transactions on Software Engineering and Methodology (TOSEM), 2013, 22(4): 31
- [11] Rothermel G, Untch R H, Chu C. Prioritizing test cases for regression testing[J]. IEEE Transactions on Software Engineering, 2001, 27(10): 929-948
- [12] Yoo S, Harman M. Regression testing minimization, selection and prioritization: a survey[J]. Software Testing, Verification and Reliability, 2012, 22(2): 67-120
- [13] Jiang B, Zhang Z, Chan W K. Adaptive random test case prioritization[C]// 24th IEEE/ACM International Conference on Automated Software Engineering, 2009(ASE'09). IEEE, 2009: 233-244
- [14] Xue X, Namin A S. How Significant is the Effect of Fault Interactions on Coverage-Based Fault Localizations? [C]// 2013 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement. IEEE, 2013: 113-122
- [15] Wang Ke-chao, Wang Tian-tian, Su Xiao-hong, et al. Test Case Selection for Improving the Effectiveness of Software Fault Localization[J]. Journal of Computer Research and Development, 2014, 51(4): 865-873 (in Chinese)
王克朝, 王甜甜, 苏小红, 等. 面向有效错误定位的测试用例优选方法[J]. 计算机研究与发展, 2014, 51(4): 865-873
- [16] Yu, Yan-bing, Jones J A, et al. An empirical study of the effects of test-suite reduction on fault localization[C]// Proceedings of the 30th International Conference on Software Engineering. ACM, 2008: 201-210

(上接第 192 页)

- [4] Johnson M, Rosebrugh R. Fibrations and universal view updatability[J]. Theoretical Computer Science, 2007, 388: 109-129
- [5] Johnson M, Rosebrugh R, Wood R J. Lenses, fibrations and universal translations[J]. Mathematics Structure in Computer Science, 2012, 22: 25-42
- [6] Tews H. Coalgebra method for object-oriented specification[D]. Dresden, Germany: Institute Theoretische Informatik, Technischen Universiy Dresden, 2002
- [7] Ghani N, Johann P, Fumex C. Generic fibrational induction[J]. Logical Methods in Computer Science, 2012, 8(2): 1-27
- [8] Miao De-cheng, Xi Jian-qing, Jia Lian-yin, et al. Formal language algebraic model [J]. Journal of South China University of Technology (Natural Science Edition), 2011, 39(10): 74-78 (in Chinese)
苗德成, 奚建清, 贾连印, 等. 一种形式语言代数模型[J]. 华南理工大学学报(自然科学版), 2011, 39(10): 74-78
- [9] Hermida C, Jacobs B. Structural induction and coinduction in a fibrational setting[J]. Information and Computation, 1998, 145(2): 107-152
- [10] Barr M, Wells C. Category theory for computing science[M]. New York: Prentice-Hall, 1990: 252-270
- [11] He Wei. Category theory[M]. Beijing: Science Press, 2006: 23-69 (in Chinese)
贺伟. 范畴论[M]. 北京: 科学出版社, 2006: 23-69
- [12] Ghani N, Johann P, Fumex C. Indexed induction and coinduction, fibrationally[J]. Logical Methods in Computer Science, 2013, 9(3-6): 1-31
- [13] Hermida C. Fibrations, Logical Predicates and Related Topics [D]. Edinburgh, UK: University of Edinburgh, 1993
- [14] Su Jin-dian, Yu Shan-shan. Coinductive data types and their applications in programming languages [J]. Computer Science, 2011, 38(11): 114-118 (in Chinese)
苏锦钿, 余珊珊. 程序语言中的共归纳数据类型及其应用[J]. 计算机科学, 2011, 38(11): 114-118
- [15] Hagino T. A categorical programming language [D]. Edinburgh, UK: Laboratory for Foundations of Computer Science, Dept of Computer Science, University of Edinburgh, 1987
- [16] Nogueira P, Moreno-Navarro J. Bialgebra views: a way for polytypic programming to cohabit with data abstract[C]// Proceedings of the ACM SIGPLAN Workshop on Generic Programming. ACM New York, NY: 2008: 61-73
- [17] Poll E. Subtyping and inheritance for categorical datatypes[J]. RIMS Lecture Notes, 1998, 1023: 112-125
- [18] Hinze R. Reasoning about codata [J]. Lecture Notes in Computer Science, 2010, 6299: 42-93
- [19] Gimenez E, Casteran P. A Tutorial on Co-inductive Types in Coq[OL]. May 1998. <http://www.labri.fr/perso/casteran/Rec-Tutorial.pdf>
- [20] Vene V. Categorical programming with inductive and coinductive types[D]. Tartu, Estonia: University of Tartu, 2000