第44卷第7期 中国科学技术大学学报 Vol. 44, No. 7
2014年7月 JOURNAL OF UNIVERSITY OF SCIENCE AND TECHNOLOGY OF CHINA Jul. 2014

文章编号:0253-2778(2014)07-0599-06

抽象解释全总域模型

王蓁蓁[1,2],倪庆剑[3],张志政[3],邢汉承[3]

(1. 金陵科技学院信息技术学院,江苏南京 211169; 2. 江苏省信息分析工程实验室,江苏南京 211169;
3. 东南大学计算机科学和工程学院,江苏南京 210096)

摘要:抽象解释自1977年提出后,许多作者做了大量工作,将抽象解释理论应用于程序分析和验证研究等领域.本文为有关抽象解释论述构造了一个统一模型,称为抽象解释的全总域模型,目前现存的有关抽象解释文献所采取的框架都相容于全总域模型,且是等价的.在此基础上,我们还提出有关抽象解释理论需要解决的几个基本问题.模型和问题都可以作为今后抽象解释理论发展的参考基点.

关键词:抽象解释;语义;全总域;完备性

中图分类号:TP311 **文献标识码**:A doi:10.3969/j.issn.0253-2778.2014.07.009

引用格式: Wang Zhenzhen, Ni Qingjian, Zhang Zhizheng, et al. Universe model of abstract interpretation[J]. Journal of University of Science and Technology of China, 2014,44(7):599-604.
王蓁蓁,倪庆剑,张志政,等. 抽象解释全总域模型[J]. 中国科学技术大学学报,2014,44(7):599-604.

Universe model of abstract interpretation

WANG Zhenzhen[1,2], NI Qingjian[3], ZHANG Zhizheng[3], XING Hancheng[3]

(1. *School of Information Technology, Jinling Institute of Technology, Nanjing* 211169, *China*;
2. *Information Analysis Engineering Laboratory of Jiangsu Province, Nanjing* 211169, *China*;
3. *School of Computer Science & Engineering, Southeast University, Nanjing* 210096, *China*)

Abstract: Since its introduction in 1977, abstract interpretation has inspired a lot of research and is now widely applied in program analyses and verification fields. Therefore, a universe model was constructed for existing studies on abstract interpretation, which unifies and is equivalent to all the current frameworks of abstract interpretation. Based on this, several fundamental problems were raised about abstract interpretation that need to be solved. This model and the relevant problems can be viewed as the basic points for further development of abstract interpretation theory.

Key words: abstract interpretation; semantics; universe domain; completeness

0 引言

自1977年提出抽象解释理论以来[1],抽象解释理论已经在很多领域得到广泛应用[1-12],其应用范围涵盖了程序变换,程序调试,程序静态分析等方法,并且有些应用,如程序静态分析已经在嵌入式系统的自动验证中取得了成功.事实上,经典抽象解释理论是在Golois connections或与之等价的closure

收稿日期:2014-03-21;**修回日期**:2014-06-15

基金项目:金陵科技学院科研基金(jit-n-201305),2013年度江苏省高校自然科学研究面上自筹经费项目(13KJD520005)资助.

作者简介:王蓁蓁(通讯作者),女,1975年生,博士后/副教授. 研究方向:软件测试,人工智能,程序分析.
E-mail: wangzhenzhen@seu.edu.cn

operators 的关于抽象论域的框架下进行的，然而如何建立一个通用的、具有较强表达能力且代价合理的抽象理论框架以处理不同应用领域中的复杂结构及特性是抽象解释理论有待进一步探讨和研究的问题[12]. 本文就是对此种范式下的抽象解释理论进行总结，为经典抽象理论构造统一模型. 该模型的主要作用是：①为经典抽象理论构造统一模型，并在此模型上讨论主要问题，例如完备性问题. ②深化经典抽象理论的概念，例如提出本质函数和亚本质函数等概念以及泛型概念的广义表述. ③该模型可以进一步启发我们把抽象解释理论应用到软件工程领域：从设计到实现的过程，即从抽象设计到具体论域的实践.

1 构造全总域模型

1.1 具体语义论域

用 S 表示一个具体系统(其中元素可以是无限的)，例如在经典模型检验里，S 是一个(具体) Kripke 结构，其中 S 中元素表示系统的状态. 以下我们用状态指称 S 中的元素，假设 S 是可数无限的，并称它为状态空间. 我们用$\wp(S)$表示 S 的幂集，用以表示系统 S 的语义论域. $\wp(S)$中的元素是 S 的子集，是 S 中状态组成的集合. 对于任意 $a\in\wp(S)$，可以是 S 中有限子集合或是 S 中无限子集合. 即

$$a=\{s_1,s_2,\cdots,s_n\}\text{ 或 }a=\{s_1,s_2,\cdots,s_n,\cdots\}.$$

其中，$s_i\in S$，是 S 中的状态，$i=1,2,\cdots,n,\cdots$. 关于 a 的意义，叙述如下：

a 可以表示一个(成员关系)谓词 P，即每一个 s_i 使该谓词 P 为真. 如果适当扩展$\wp(S)$和/或 S(扩展后仍记为$\wp(S)$)，则 a 还可以表示一个(任意元)的谓词.

a 可以表示某一属性(性质)φ，即每一个 s_i 具有该属性(或该性质)φ. 为了简单叙述下面引进的语义操作算子的单调性，在考虑属性 φ 时，我们限制“否定性质”到基本原子层次上，这与(模型检验)讨论 LTL 逻辑公式时所作的限制一样. 可以作些细致处理，像文献[1]那样，引进符号 $p::=+\mid-$. 定义 $\leqslant^{+}\triangleq\leqslant$，$\leqslant^{-}\triangleq\geqslant$. 于是在讨论“序”时，增加条件“$\leqslant^{p}$”，则可以统一各种单调性到同一个方向上. 为了简单起见，本文不这样做.

a 可以表示状态空间 S 的某些轨道可达到的状态集合，例如在考虑程序“不变式”属性时，我们常常是这样做的(见文献[1]). 当然，a 也可以表示状态空间 S 里一条(具有某性质)轨道，即 $s_1,s_2,\cdots$ 是从 s_1 出发的一条轨道的顺序出现的状态集合. 不过为了表示循环轨道，我们必须扩展语义域$\wp(S)$和/或 S，仍记为$\wp(S)$，则 a 也可以表示状态空间里的循环轨道.

a 可以表示一个类型 T，准确地说它是类型 T 的值域，即 a 中所有元素 s_i 都属于类型 T.

总之，我们在非常广泛的意义上，理解$\wp(S)$中的元素的语义特征(如果需要的话，还可以扩展 S 和/或$\wp(S)$使它们可以表达更多的内容，例如文献[1]的 Temporal models)，为了确定起见，我们指称$\wp(S)$中的元素为属性，这样也较直观. 在集合包含的自然序下，$\wp(S)$任一子集都有最小上界和最大下界，所以$\wp(S)$是一个完备格. 若 $a,b\in\wp(S)$，$a\sqsubseteq b$，则认为 a 较 b“精确”，即 b 是 a 的近似.

1.2 抽象语义总域

在经典抽象解释框架里，依据具体状态幂集的抽象解释，Golois 连接和闭包操作是等价的. 闭包操作具有独立于抽象元素表示法的优点，如果推理不涉及抽象表示内涵时，用闭包操作作为抽象语义领域是比较方便的. 因此我们用 $\mathrm{uco}(\wp(S))$表示系统 S 的抽象语义全总域. 其中，任意元素 $\mu\in\mathrm{uco}(\wp(S))$是一个闭包，即 μ 是具体语义域$\wp(S)$的一个抽象语义域，而全总域 $\mathrm{uco}(\wp(S))$是所有抽象语义域的集合. 对于闭包 $\mu\in\mathrm{uco}(\wp(S))$，它可以从两个等价角度定义：

μ 是$\wp(S)$上单调、幂等、扩展的操作，因此可以把 μ 看作是函数.

即设 a，b 是$\wp(S)$里任意两个元素：

若 $a\sqsubseteq b$，则 $\mu(a)\sqsubseteq\mu(b)$. 单调性；

$\mu\circ\mu(a)\triangleq\mu(\mu(a))=\mu(a)$. 幂等性；

$\mu(a)\sqsupseteq a$. 扩展性.

μ 是由它的不动点唯一确定的，且 μ 在$\wp(S)$上操作的所有不动点组成的集合与 μ 的像集$\mu(\wp(S))$一致. 因此可以把 μ 看成是$\wp(S)$中的子集合. 即

$$\mu(\wp(S))=\{a\in\wp(S)\mid\mu(a)=a\}.$$

今后指称任一闭包 $\mu\in\mathrm{uco}(\wp(S))$，我们互用它的函数和集合两种表示. 值得注意的是，$a\in\wp(S)$表示系统 S 某一属性 φ，于是由 $\mu\in\mathrm{uco}(\wp(S))$的扩展性 $\mu(a)\sqsupseteq a$，就可以认为一个闭包是将一个任意

的 $a\in\wp(S)$到 a 的满足某个性质的最小超集的映射,而那个最小超集可以看作是属性 φ 的从上面的一个最为"精确"的近似.实际上,关于 μ 的两种等价定义,我们有:$\forall\, a\in\wp(S)$,$\forall\,\mu\in \mathrm{uco}(\wp(S))$,

$$\mu(a)=\bigcap\{b\mid a\subseteq b\ \text{且}\ b\in\mu\}.$$

所以,若 $\mu(a)=b'$,则 b'就是最接近 a 的一个近似表示.

在经典抽象解释框架下,这种近似就称为抽象.这也是我们称任一闭包 $\mu\in\mathrm{uco}(\wp(S))$为具体语义域$\wp(S)$的一个抽象语义域的原因,而 $\mathrm{uco}(\wp(S))$就是所有抽象语义域的全总域.

对于任一闭包 $\mu\in\mathrm{uco}(\wp(S))$,$\mu$ 由 S 中的子集组成,按照集合的包含自然序,μ 可以构成一个完备格,虽然它未必是$\wp(S)$上的完备子格.若 $a,b\in\mu$,$a\subseteq b$,则称 a 较 b 精确,即 b 是 a 的一个近似.对于全总域 $\mathrm{uco}(\wp(S))$,我们在抽象域之间规定一个序"$\sqsubseteq$",它就是通用的函数之间逐点序.这样任意两个闭包 $\rho,\eta\in\mathrm{uco}(\wp(S))$,$\rho\sqsubseteq\eta$,即 $\forall\, a\in\wp(S)$,都有 $\rho(a)\subseteq\eta(a)$(等价地 $\eta(\wp(S))\subseteq\rho(\wp(S))$),它表示在 $\mathrm{uco}(\wp(S))$全总域,ρ 较 η 精确些,η 较 ρ 抽象.于是,$\mathrm{uco}(\wp(S))$在逐点序"$\sqsubseteq$"下也构成一个完备格.

1.3　具体语义操作和抽象语义操作

我们用 f 表示$\wp(S)$具体语义域上的一个语义操作,为了简单起见,只考虑

$$f:\wp(S)\to\wp(S)$$

是$\wp(S)$上的一元单调函数情形.因为对一般情况的单调函数可以类似处理.例如,当我们考虑完备性问题时,文献[2]命题 3.5Ⅷ指出,可以把 n 元函数的完备性问题化归为一组一元函数的完备性问题.更复杂一些,上文提及扩充 $\wp(S)$ 和/或 S,可以使$\wp(S)$具有我们需要的语义属性,同样我们能使用递归配对函数或者是参数分离来处理多元函数,甚至利用并行系统还可以处理不确定的语义操作,因此只考虑一元单调函数并不丧失一般性.下面我们互用语义操作和语义函数两个术语.

对于任意语义函数 f,若 $\rho,\eta\in\mathrm{uco}(\wp(S))$是两个任意闭包,令

$$f^{\rho,\eta}\triangleq\eta\circ f\circ\rho \tag{1}$$

为 f 的关于$<\rho,\eta>$的抽象语义操作.在经典抽象解释框架里,$f^{\rho,\eta}$为 f 在抽象域 ρ,η 上最正确近似抽象语义操作.ρ 和 η 也可以是同一个闭包.

合理性和完备性是经典抽象解释理论里最基本的两个概念(见文献[2-4]).

f 是任意一个具体语义函数,ρ,η 是任意两个闭包,$f^{\sharp}:\rho\to\eta$ 是对应的(在$<\rho,\eta>$)上的抽象语义函数,若

$$\eta\circ f\sqsubseteq f^{\sharp}\circ\rho \tag{2}$$

则称 $f^{\sharp}$ 是合理抽象语义函数,合理性与近似性等价.由此,最正确近似 $f^{\rho,\eta}$自动满足式(2),这是由于 ρ,η 都是单调函数且满足幂等性的缘故.若 $f^{\sharp}$ 是$<\rho,\eta>$上任意一个合理抽象函数,因为 $\eta\circ f\sqsubseteq f^{\sharp}\circ\rho$,于是 $\eta\circ f\circ\rho\sqsubseteq f^{\sharp}\circ\rho\circ\rho$,根据 ρ 的幂等性,可得 $\eta\circ f\circ\rho\sqsubseteq f^{\sharp}\circ\rho$,再根据 $f^{\rho,\eta}$的定义和 $f^{\sharp}$ 的定义域,所以 $f^{\rho,\eta}\sqsubseteq f^{\sharp}$ 关系成立,这也是我们称 $f^{\rho,\eta}$是最正确的近似的原因.

沿用上面的术语,如果等式(2)成立,即

$$\eta\circ f=f^{\sharp}\circ\rho \tag{3}$$

则称抽象解释 $f^{\sharp}$ 关于 f 在$<\rho,\eta>$上是完备的.下面我们将证明这时 $f^{\rho,\eta}$关于 f 在$<\rho,\eta>$上也是完备的.所以为了叙述方便,这时也称$<\rho,\eta>$关于 f 是完备的.如果从上下文不会混淆,我们互用 $f^{\sharp}$($f^{\rho,\eta}$)完备或$<\rho,\eta>$完备两个术语.

更进一步,如果给定$<\rho,\eta>$,$f^{\sharp}$ 关于 f 是完备的,则 $f^{\rho,\eta}$也是完备的,并且 $f^{\rho,\eta}$与 $f^{\sharp}$ 一致.实际上:

$f^{\sharp}=$(因为 ρ,η的幂等性);

$f^{\sharp}\circ\rho=$(由式(3)$f^{\sharp}$ 完备性定义);

$\eta\circ f\sqsubseteq$(由式(2)$f^{\rho,\eta}$的合理性);

$\eta\circ f\circ\rho\sqsubseteq$(由近似性:$f^{\rho,\eta}\sqsubseteq f^{\sharp}$);

$f^{\sharp}$.

鉴于此,式(3)也可以改写为:

$$\eta\circ f=\eta\circ f\circ\rho \tag{3a}$$

本文只讨论最正确近似抽象 $f^{\rho,\eta}$,除非特殊声明,我们把 $f^{\rho,\eta}$简记为 $f^{\sharp}$,即在定义抽象语义操作时,只用式(1).因为 $f^{\sharp}=f^{\rho,\eta}=\eta\circ f\circ\rho$,它是 $\rho\to\eta$ 上相应于 f 的抽象语义函数,所以直观上它是这样的抽象语义函数:首先强制它的实参 $x\in\wp(S)$为闭包 ρ 中的元素,接着运用 f,其后强制结果为闭包 η 中的元素.不难看出,若 $x\in\rho$,则 $f^{\rho,\eta}(x)=\eta\circ f\circ\rho(x)\in\eta$,因此,$f^{\rho,\eta}$实质上是一个特殊的"具体"语义函数,限制它的定义域和值域分别到 ρ 闭包集合和 η 闭包集合上.即 $f^{\rho,\eta}$是 $\rho\to\eta$ 函数,是在$<\rho,\eta>$上关于 f 的抽象语义操作.

令 $\mathcal{F}=\{f\mid f:\wp(S)\to\wp(S)$,(单调)具体语义函

数},是所有具体语义函数的集合,用 $F\subset\mathscr{F}$ 表示 $\mathscr{F}$ 的子集合,它是若干个语义操作的集聚.

用 $\mathscr{F}^{\sharp}=\{f^{\sharp}\mid \exists f\in\mathscr{F},\exists \rho,\eta\in \mathrm{uco}(\wp(S))$,使得 $f^{\sharp}=\eta\circ f\circ\rho\}$表示所有抽象语义函数的集合,同样记 $F^{\sharp}\subset\mathscr{F}^{\sharp}$ 是相应于具体语义函数集聚 F 的抽象语义函数集聚.

1.4 全总域模型

按照前面的叙述,实际上我们已经为抽象解释框架构造了一个模型,称之为全总域模型,它是一个五元组:

$$\mathscr{A}=\{S,\wp(S),\mathrm{uco}(\wp(S)),\mathscr{F},\mathscr{F}^{\sharp}\}\qquad(4)$$

式中,每个符号的解释如上文所述.

从模型的构造流程可以看出,关于经典抽象解释所采用的框架,在隐蔽形式下或者在等价形式下,都与式(4)的框架相容.利用式(4)模型,我们至少可以讨论有关抽象解释大部分重要问题.

2 理论性问题

2.1 最优完备存在性

在式(4)框架里,对任意具体语义函数 f,无论 η 如何确定,只要令 $\rho=\wp(S)$,则式(3a)成立,即

$$\eta\circ f=\eta\circ f\circ\rho.$$

即,只要选择 $\rho=\wp(S)$,对任意的 f,任意的 η,抽象语义函数 $f^{\rho,\eta}$关于 f 是平凡的完备的.

同样,若选择 $\eta=\{S\}$,即 η映射$\wp(S)$任一子集为基本集合 S,这时对任意 $f\in\mathscr{F}$,任意 $\rho\in\mathrm{uco}(\wp(S))$,式(3a)也成立.

$$\eta\circ f=\eta\circ f\circ\rho=\{S\}.$$

即,只要选择 $\eta=\{S\}$,$\forall f\in\mathscr{F}$,$\forall\rho\in\mathrm{uco}(\wp(S))$,抽象语义函数 $f^{\rho,\eta}$关于 f 是平凡完备的.

根据文献[2](命题 3.5ⅷ),若$<\rho,\eta>$关于 f 完备,$\rho\sqsupseteq\delta\in\mathrm{uco}(\wp(S))$和 $\eta\sqsubseteq\beta\in\mathrm{uco}(\wp(S))$,则$<\delta,\beta>$也关于 f 完备.即,若$<\rho,\eta>$关于 f 完备,将 ρ"精确化"到 $\delta(\delta(\wp(S))\supseteq\rho(\wp(S)))$,将 η"抽象化"到 $\beta(\beta(\wp(S))\subseteq\eta(\wp(S)))$,则$<\delta,\beta>$也关于 f 完备.δ和β是从两个不同方向分别对 ρ 和 η 的"改良",$\rho\rightarrowtail\delta$是延展,$\eta\rightarrowtail\beta$是简化.

鉴于上述考虑,我们提出最优完备性概念.

定义 2.1 最优完备.设$<\rho,\eta>$关于 f 完备,如果对任意关于 f 的完备抽象$<\rho',\eta'>$,都有:

$$\rho\sqsupseteq\rho',\ \eta\sqsubseteq\eta'\qquad(5)$$

则称 $\rho\circ f\circ\eta$ 是 f 的最优完备抽象语义函数,也称$<\rho,\eta>$是 f 的最优完备抽象语义域.

根据文献[2]命题 3.5ⅷ,对于最优完备$<\rho,\eta>$而言,无需进一步分别对 ρ在精确化方向,对 η在抽象化方向加以改进.同样,也无法进一步分别对 ρ在抽象化方向,对 η在精确化方向加以改进,若存在这样的改进$<\delta,\beta>$,由于 δ是 ρ 的抽象化,即 $\rho\sqsubseteq\delta$,由最优完备定义,$\rho\sqsupseteq\delta$,所以 $\rho=\delta$;由于 β是 η 的精确化,即 $\beta\sqsubseteq\eta$,由最优完备定义,$\eta\sqsubseteq\beta$,所以 $\beta=\eta$.

因此,我们提出如下问题.

问题 1 对于具体语义函数 f,是否存在它的最优完备抽象语义函数 $f^{\sharp}$,或者它的最优完备抽象语义域?

对于函数族 F,若$<\rho,\eta>$对于 F 中每一个函数都完备,则称$<\rho,\eta>$关于 F 完备.同样我们也可以对于函数族 F 定义最优完备问题,即若存在$<\rho,\eta>$关于 F 完备,且对任意关于 F 的完备$<\rho',\eta'>$而言,条件(5)成立,则称$<\rho,\eta>$是 F 的最优完备抽象语义域.

问题 2 对于语义函数族 $F=\{f_1,f_2,\cdots,f_n\}$是否存在它的最优完备抽象语义域?如果$\forall i=1,2,\cdots,n$,$<\rho_i,\eta_i>$关于 f_i 最优完备,那么 F 的最优完备抽象语义域(如果存在的话)与$<\rho_i,\eta_i>$之间有什么样关系?即能从$<\rho_i,\eta_i>i=1,2,\cdots,n$构造 F 的最优完备?

根据文献[2]的定义 5.1,设 F 是具体语义函数族,$\rho,\eta\in\mathrm{uco}(\wp(S))$,定义算子 C_F^{ρ},S_F^{η} 如下:

$$C_F^{\rho}:\mathrm{uco}(\wp(S))\to\mathrm{uco}(\wp(S))$$

其中,$C_F^{\rho}(\eta)\triangleq\bigsqcap\{\beta\in\mathrm{uco}(\wp(S))\mid\eta\sqsubseteq\beta,<\rho,\beta>$关于 F 完备$\}$.

$$S_F^{\eta}:\mathrm{uco}(\wp(S))\to\mathrm{uco}(\wp(S))$$

其中,$S_F^{\eta}(\rho)\triangleq\bigsqcup\{\delta\in\mathrm{uco}(\wp(S))\mid\delta\sqsubseteq\rho,<\rho,\eta>$关于 F 完备$\}$.

再根据文献[2]的定义 5.2,设 F 是具体语义函数族,$\rho,\eta\in\mathrm{uco}(\wp(S))$.

(ⅰ)若$<\rho,C_F^{\rho}(\eta)>$关于 F 完备,则称 $C_F^{\rho}(\eta)$是关于 F 相对于 ρ 的 η 完备核.

(ⅱ)若$<S_F^{\eta}(\rho),\eta>$关于 F 完备,则称 $S_F^{\eta}(\rho)$是关于 F 相对于 η 的 ρ 的完备壳.

根据 $S_F^{\eta}(\rho)$的定义,可知 $S_F^{\eta}(\rho)$是一切满足定义的 $\delta(\delta\sqsubseteq\rho)$的上确界,于是 $S_F^{\eta}(\rho)\sqsubseteq\rho$,若它是(相对于 η关于F)ρ完备壳,则$<S_F^{\eta}(\rho),\eta>$是$<\rho,\eta>$在 ρ

上向精确化方向的一次改良.

同样,$C^{\circ}_{F}(\eta)\sqsupseteq\eta$,若它是(相对于 ρ 关于 F)η 的完备核,则$<\rho,C^{\circ}_{F}(\eta)>$是$<\rho,\eta>$在 η 上向抽象化方向的一次改良.

问题 3　给定具体语义函数族 F,其最优完备(如果存在)$<\rho,\eta>$能否通过初始化 $\rho=\{S\}$,$\eta=\wp(S)$递归运用 $S^{\sharp}_{F}(\rho)$、$C^{\circ}_{F}(\eta)$算子分别对 ρ 在精确化方向、对 η 在抽象化方向进行改良而得?

2.2　最优函数存在性

根据 2.1 节关于抽象语义函数 $f^{\sharp}=\eta\circ f\circ\rho$ 的解释,对于给定的闭包 $\rho,\eta\in uco(\wp(S))$,仿照 λ 演算,定义(其中 $f:\mathscr{F}$,表示 $f\in\mathscr{F}$):

$$\rho\rightarrow\eta\triangleq\lambda f:\mathscr{F}.\ \eta\circ f\circ\rho \tag{6}$$

其直观意义是所有的关于闭包$<\rho,\eta>$上的抽象语义函数的集合.现在我们证明 $\rho\rightarrow\eta$ 的元素是外延的.

$\rho\rightarrow\eta$的任务是取任意具体语义函数 f,并产生这样的函数 $\eta\circ f\circ\rho$,即它把每一个函数 f 强制到从 ρ 的值域到 η 的值域的抽象函数,因此,若对所有 $x\in\rho$,我们有:

$$(\eta\circ f_1\circ\rho)x=(\eta\circ f_2\circ\rho)x.$$

注意到限制 $x\in\rho$ 无关紧要,所以 $\forall x\in\wp(S)$,都有:

$$(\eta\circ f_1\circ\rho)x=(\eta\circ f_2\circ\rho)x.$$

于是,$\lambda x.(\eta\circ f_1\circ\rho)x=\lambda x.(\eta\circ f_2\circ\rho)x$.

根据复合的意义,结合上式,应该有 $\eta\circ f_1\circ\rho=\eta\circ f_2\circ\rho$,这意味着 $\rho\rightarrow\eta$ 的值域对于每个在$\wp(S)$上入射 ρ 到 η 的单调函数恰好包含一个代表,这说明 $\rho\rightarrow\eta$模型是外延的.下面的讨论就是基于上面叙述的直观意义,把 $\rho\rightarrow\eta$ 看作关于闭包$<\rho,\eta>$上的抽象语义函数的集合.

不难看出,若 f 已经是一个从 ρ 到 η 的(具体)语义函数,即一旦 $x\in\rho$ 就有 $f(x)\in\eta$,那么对所有的 $x\in\rho$,有 $f^{\sharp}(x)=\eta\circ f\circ\rho(x)=f(x)$.基于这样的观察,我们引入下述概念.

定义 2.2　本质函数和亚本质函数.

① 若 $\forall x\in\rho,f(x)=f^{\sharp}(x)\in\eta$,则称 f 是$\rho\rightarrow\eta$里本质(具体)语义函数.换言之,把 f 的定义域限制到 ρ 闭包上,记之为 $f_{\downarrow\rho}$,则 $f_{\downarrow\rho}\equiv f^{\sharp}$.

② 若 $\forall x\in\wp((S),\eta\circ f=f^{\sharp}=\eta\circ f\circ\rho$,则称 f 是$\rho\rightarrow\eta$亚本质(具体)语义函数.对于亚本质函数 f,$<\rho,\eta>$是完备的抽象语义域.

亚本质(本质)函数是 $\rho\rightarrow\eta$类中的特殊函数族.直观意义是,当设计一个系统(检验一个系统也一样)时,我们对系统的特性、功能和行为都有一个预先的设想,这些设想就是抽象语义域及其在抽象域上的操作,当我们实现系统时,具体的语义域及其上的操作受到技术及其环境的限制,自然要复杂迂回得多,所以本质函数和亚本质函数反映了理想操作的忠实可靠的实现.在这个意义上,它们都是最优函数族.

问题 4　如果 ρ,η 抽象语义域构想完美,能否解决技术上的问题,使得实现 ρ,η 的具体语义域上的操作都是本质的或亚本质函数?或者说,能否改进具体语义操作到至多是只用亚本质函数就能解决实际(例如完备性)问题?

一般意义上,$<\rho,\eta>$抽象域比具体语义域$\wp(S)$简单,而 $\rho\rightarrow\eta$ 上的抽象语义函数也比$\wp(S)\rightarrow\wp(S)$上具体语义函数少,所以在模型检验应用时,我们从两个方面(域和操作)作了简化,只要解决了完备性问题,那么模型检验中高期望强保留特性就能满足.可惜的是,从问题 1 到问题 4,至今仍然没有"彻底"解决.文献[2]指出:以可能的最好方式,即最小化延展或简化正在考虑的抽象域和操作算子,使得抽象解释完备,仍然是一个开放问题.甚至,寻找某些合理的很强的条件施加于具体语义域和/或具体操作算子,以致使抽象域的不动点完备壳存在,一般来说都是不能保证的.因此,对于上述问题,是否能够定义某种"近似"标准和"近似"计算,寻找相关问题的某种意义上的"最好"解决,这些都是值得探讨的.

2.3　与不可预知多态模型类比

简写 $\mathscr{V}=uco(\wp(S))$,并用记号 $\rho:\mathscr{V}$ 表示 $\rho\in uco(\wp(S))$,定义:

$$\rightarrow^{\sharp}\triangleq\lambda\rho:\mathscr{V}.\lambda\eta:\mathscr{V}.(\lambda f:\mathscr{F}.\ \eta\circ f\circ\rho) \tag{7}$$

表示所有的抽象函数的集合,式(7)表明它可以看作式(4)全总域模型 $\mathscr{A}$ 里的元素 $\mathscr{F}^{\sharp}$,即$\rightarrow^{\sharp}=\mathscr{F}^{\sharp}$.改写$\rightarrow^{\sharp}$为 $\forall$,即

$$\forall=\lambda f:\mathscr{F}.(\lambda\rho:\mathscr{V}.\lambda\eta:\mathscr{V}.\ \eta\circ f\circ\rho)=\lambda f:\mathscr{F}.(\lambda<\rho,\eta>:\mathscr{V}\times\mathscr{V}.\ \eta\circ f\circ\rho) \tag{8}$$

于是 $\forall f$ 的定义域为 $\mathscr{V}\times\mathscr{V}=uco(\wp(S))\times uco(\wp(S))$,值域是$\{\eta\circ f\circ\rho\}$,即

$$\forall f=\lambda<\rho,\eta>:\mathscr{V}\times\mathscr{V}.\ \eta\circ f\circ\rho:\mathscr{V}\times\mathscr{V}\rightarrow\{\eta\circ f\circ\rho\} \tag{9}$$

直观上,$\forall f$ 是抽象函数族,它是具体操作 f 的泛化,在不同的$<\rho,\eta>$上,呈现出不同的特性,具有"多

态”特征.在某种意义上,它体现了多态性或泛型性.

现在扩展式(4)的全总域模型 $\mathscr{A}$ 如下,扩展后记为 $\mathscr{A}'$.

$$\mathscr{A}'=\{S,\wp(S),\mathrm{uco}(\wp(S)),\mathscr{F},\mathscr{F}^{\sharp},\{\rho\rightarrowtail\eta\},\{\forall f\}\} \tag{10}$$

是 7 元组,其中符号是自明的.例如,$\{\rho\rightarrowtail\eta\}$ 是一切形如式(6)定义的函数族的集聚,其中 $\rho\rightarrowtail\eta$ 是在固定的闭包对$<\rho,\eta>$上讨论每一个具体操作相应的抽象操作,重点是具体操作和抽象操作之间的对应关系;$\{\forall f\}$ 是一切形如式(9)定义的函数族的集聚,其中 $\forall f$ 是讨论具体操作 f 的在每一对$<\rho,\eta>$闭包上的多态表现,即 $\forall f$ 是 f 的泛型.

式(10)模型总的想法是:一个多态函数是一种可以作用在多种“类型”实参上的函数.在某种意义上,该多态函数在各个类型上使用了“本质同样的算法”.例如,对于具体语义函数 f,式(9)定义的“$\forall f$”,它可以作用在所有“类型”$<\rho,\eta>$上.因为在每个$<\rho,\eta>$上,它都是相应于 f 的抽象语义函数 $\eta^{\circ}f^{\circ}\rho$,所以它们的算法在本质上具有“共性”.关于它的具体实例可以参阅文献[5].

为了能够从大量数据中寻找有意义的新的关系和模式,数据挖掘技术从而诞生.根据大量经验数据,通过各种定性分析或者定量分析,对数据加以整理,归纳产生抽象解释,进一步将这种抽象解释总结成理论,甚至上升到“定律”高度.如果从这个角度考虑问题,模型(4)和(10)还具有方法论意义.

前面提出的问题 1~3,都是立足在$\{\rho\rightarrowtail\eta\}$上讨论,怎样改良$<\rho,\eta>$使得具体语义函数和抽象语义函数之间满足完备性关系,而问题 4 是立足于$\{\forall f\}$,讨论怎样设计良好操作,使它具有“安全”的泛型,即它一切泛化都具有完备特性.再把 $\mathscr{A}'$ 与文献[5]相比较,$\mathscr{A}'$ 和文献[5]使用的 $\mathscr{P}\omega$ 的闭包构造出来的模型在许多方面都是类似的.实际上,$\mathscr{A}$ 和 $\mathscr{A}'$ 模型都是受到文献[5]的启发构造出来.

问题 5 基于 $\mathscr{A}'$,我们能否构造一个有关抽象解释计算理论?

3 结论

通过本文的研究,我们针对现有的抽象解释理论构造了一个统一模型,称为抽象解释全总域模型.在该模型上,可以很方便地讨论经典抽象理论的一些问题,如完备性问题等.并且在该模型的基础上,我们提出了本质函数和亚本质函数的概念以及对泛型概念给出了其广义的表达方式.这个模型一方面可以深化我们对经典抽象理论概念的理解,另一方面可以作为抽象解释理论的基础.

参考文献(References)

[1] Cousot P, Cousot, R. Temporal abstract interpretation [C]// Proceedings of the 27th ACM Symposium on Principles of Programming Languages. Boston, USA: ACM Press, 2000: 12-25.

[2] Giacobazzi R, Ranzato F, Scozzari F. Making abstract interpretations complete [J]. Journal of the ACM 2000, 47(2): 361-416.

[3] Cousot P, Cousot R. Systematic design of program analysis frameworks [C] // Proceedings of the 6th ACM Symposium on Principles of Programming Languages. San Antonio, USA: ACM Press, 1979: 269-282.

[4] Cousot P, Cousot R. Abstract interpretation frameworks [J]. Journal of Logic and Computation, 1992, 2(4): 511-547.

[5] Mitchell J C. 程序设计语言理论基础[M]. 许满武, 徐建, 袁宜, 等译, 北京: 电子工业出版社, 2006.

[6] Cousot P, Cousot R. Abstract interpretation: A unified lattice model for static analysis of programs by construction of approximation of fixpoints [C] // Proceedings of the 6th Annual ACM Symposium on Principles of Programming Languages. Los Angeles, USA: ACM Press, 1977: 238-252.

[7] Ranzato F, Tapparo F. Generalized strong preservation by abstract interpretation [J]. Journal of Logic and Computation, 2007, 17(1): 157-197.

[8] Ranzato F, Tapparo F. Strong preservation of temporal fixpoint-based operators by abstract interpretation [C]// Proceedings of the 7th International Conference on Verification, Model Checking, and Abstract Interpretation. Charleston, USA: Springer, 2006: 332-347.

[9] Mayer W, Stumptner M. Abstract interpretation of programs for model-based debugging [C]// Proceedings of the 20th International Joint Conference on Artificial Intelligence. San Francisco, USA: Morgan Kaufmann Publishers, 2007: 471-476.

[10] 高鹰, 陈意云. 基于抽象解释的代码迷惑有效性比较框架[J]. 计算机学报, 2007, 30(5): 806-814.

[11] 杨波, 张明义, 谢刚. 抽象解释理论框架及其应用[J]. 计算机工程与应用, 2010, 46(8): 16-20.

[12] 李梦君, 李舟军, 陈火旺. 抽象解释理论的程序验证技术[J]. 软件学报, 2008, 19(1): 17-26.