

项重写系统等价性的归纳证明

Equivalence Proving of Term Rewriting Systems by Induction

冯 速

(北京师范大学计算机系 北京 100875)

Abstract This paper proposes a simple but powerful method for proving the equivalence in a restricted domain of two term rewriting systems based on structural and cover set induction. Comparing with the well-studied inductionless induction methods and/or proof by transformation, the method is more natural and it is suitable for automated formal proving.

Keywords Equivalence, Induction, Term rewriting systems, Automated proving

尽管学者们在计算机软件理论及相关数学理论方面做出了不懈的努力,但伴随着计算机硬件的高速发展而来的软件危机却日益严峻,其原因复杂多样,其中,主要原因之一是缺乏程序验证的方法和工具。程序设计的主要步骤有:描述问题、设计程序、实现程序及测试程序。需要注意的是,这里是测试程序的正确性而非证明程序的正确性,这样程序的正确性就不能从根本上得到保证,其原因在于现在还没有被广泛认可的程序正确性证明方法,因此,迫切需要加强这方面的研究。

一个程序是正确的是指该程序按程序设计者的意图运行,即对于每组合理输入均产生问题描述所预期的输出。从这一意义上看,问题描述是程序验证的出发点。一个正确的程序必须与问题描述匹配,即程序必须与问题描述等价。

一个项重写系统由一组定向等式组成,其简洁性及计算能力使其受到广泛的研究和应用;项重写系统为抽象数据类型提供模型、为函数型语言提供操作语义、为定理自动证明提供推理工具。在基于项重写系统的等价性证明方面也有广泛的研究^[1,5,7]。其中,大多数研究基于 Knuth-Bendix 的完备化算法^[6],即非归纳归纳法(inductionless induction)等价变换。研究取得了可喜的成果,但这些方法通常比较复杂且难以应用于等价性的形式自动证明。

本文提出一种证明两个项重写系统在特定领域等价的简单有效的归纳证明方法。与被广泛研究的非归纳归纳法及等价转换证明法相比,本方法更自然有效且适用于等价性的形式自动证明。例如,可以运用“动态项重写演算”(dynamic term rewriting calculus)及 Feng 等^[2,3]所示证明技术形式地进行等价性自动证

明。本方法可应用于程序正确性自动证明领域。

1 项重写系统的等价性

以下假定读者熟悉项重写系统的基本概念。有关术语及概念请参照文[4]及文[7]。这里仅给出文中所用的一些术语:

记 F 是函数符号的集合。 $C (\neq \emptyset)$ 是 F 的子集, C 的元素称为构造子。 X 是变量集合。对于任意的函数符号集 H , $\mathcal{F}(H, X)$ 是由 H 和 X 中符号构成的所有项的集合, $\mathcal{F}(H)$ 是由 H 中符号构成的所有基项的集合。 $\mathcal{F}(C, X)$ 的元素称为构造子项。对于任意的项 t , $\mathcal{O}(t)$ 是 t 的位集(the set of occurrences), $\leq$ 是 $\mathcal{O}(t)$ 上的字首顺序。对于 $u \in \mathcal{O}(t)$, t/u 是 t 在位置 u 的子项。对于 $u_1, \dots, u_n \in \mathcal{O}(t)$ 及 $t_1, \dots, t_n \in \mathcal{F}(F, X)$, $t[u_1 \leftarrow t_1, \dots, u_n \leftarrow t_n]$ 是用 $t_1, \dots, t_n$ 同时分别替换 t 中在位置 $u_1, \dots, u_n$ 的子项而得到的项。 σ 表示代换, $t\sigma$ 是将 σ 应用于 t 所得到的项。 R, R_1 及 R_2 表示 $\mathcal{F}(F, X)$ 上的项重写系统。 NF_R 是 R 下的正规形(normal forms)的集合。对于项 t , $NF_R(t)$ 是 R 下的 t 的正规形的集合, X_t 是在 t 中出现的变量的集合。

定义 1 设 R_1 与 R_2 为 $\mathcal{F}(F, X)$ 上任意项重写系统, $\mathcal{S}$ 为满足 $\mathcal{F}(C) \subseteq \mathcal{S} \subseteq \mathcal{F}(F)$ 的任意基项集。则 R_1 与 R_2 关于 $\mathcal{S}$ 等价, 记做 $R_1 \equiv_{\mathcal{S}} R_2$, 当且仅当 $\forall t \in \mathcal{S} \quad NF_{R_1}(t) = NF_{R_2}(t)$ 。

以上给出的等价性定义具有相当的一般性。正由于其一般性,它不能直接用于等价性归纳证明。下面将给出 R_1 与 R_2 等价的充分条件。其目的是建立适合于程序验证的等价性自动证明的形式方法,作为程序,我们通常只考虑具有(基)合流性及弱(基)终止性的项重写系统,这些性质是关于项重写系统本身的性质,而非

关于等价性的性质。多数等价性证明方法需要这些性质,这些性质本身也被广泛研究。可以通过良序证明终止性,对于特定形式的项重写系统,如具有左线性及终止性的项重写系统,可以利用重叠引理证明合流性。(基)合流性及弱(基)终止性以外的条件源于自动证明的目的,因此,如后文所述,这些条件适用于归纳证明。

2 等价性归纳证明

为了讨论适合于归纳证明的等价性充分条件,我们给出以下定义及命题。

定义 2 对于任意的项 t 和 s , $t \approx s$ 当且仅当

$$\exists \sigma_1, \sigma_2 t = s \sigma_1 \wedge s = t \sigma_2$$

即: $t \approx s$ 意味着在不考虑变量名字的情况下 t 和 s 相同。

定义 3 对于任意的项 t , t 中出现的 t 的极大构造子子项的信集 $\mathcal{O}\mathcal{A}(t)$ 是满足以下条件的集合:

$$1) \forall u \in \mathcal{O}\mathcal{A}(t) t/u \in \mathcal{F}(C, X);$$

$$2) \text{对任意的 } u \in \mathcal{O}\mathcal{A}(t), \text{如 } t/u \in \mathcal{F}(C, X) \text{ 则}$$

$$\exists u' \in \mathcal{O}\mathcal{A}(t) u' \leq u;$$

$$3) \forall u, u' \in \mathcal{O}\mathcal{A}(t) u \not\leq u'.$$

定义 4 设对项 t 有 $\mathcal{O}\mathcal{A}(t) = \{u_1, \dots, u_n\}$ 。定义 t_Δ 如下:

$$t_\Delta = t[u_1 \leftarrow x_1] \dots [u_n \leftarrow x_n]$$

其中, $x_1, x_2, \dots, x_n \in X$ 且对 $1 \leq i, j \leq n$, 当且仅当 $t/u_i = t/u_j$ 时 x_i 与 x_j 为相同变量。

t_Δ 是通过用适当的变量替换 t 中极大构造子子项而得到的项。

命题 1 对于任意项 t , 忽略变量名时 $\mathcal{O}\mathcal{A}(t)$ 及 t_Δ 唯一确定。

证明: 从定义显然。

定义 5 设 R_1 和 R_2 为 $\mathcal{F}(F, X)$ 上的项重写系统, G 为满足 $C \subseteq G \subseteq F$ 的集合。我们做以下定义:

$$1) \mathcal{D}_0 = \{t \mid \exists i \rightarrow t \in R_i, t \in \mathcal{F}(G, X)\};$$

$$2) \mathcal{D} \text{ 是 } \mathcal{D}_0 \text{ 的满足以下条件的极小子集:}$$

$$\forall t \in \mathcal{D}_0 - X \exists s \in \mathcal{D} s \rightarrow t;$$

也就是说, 首先, 我们从 R_1 和 R_2 的重写规则的左部中收集属于 $\mathcal{F}(G, X)$ 的项; 然后, 用适当的变量替换收集来的项中的所有的极大构造子子项。这构成 $\mathcal{D}_0$ 。通过删除变量项及合并相似项, 我们得到集合 $\mathcal{D}$ 。

显然, 忽略变量名时 $\mathcal{D}_0$ 及 $\mathcal{D}$ 唯一确定。 $\mathcal{D}$ 的元素称为证明项。证明项将被用于等价性的归纳证明。

以下设 G 为满足 $C \subseteq G \subseteq F$ 的任意函数符号集。

定义 6 设 $t \in \mathcal{F}(G)$ 。

1) $\mathcal{R}(t)$ 是关于 $R_1 \cup R_2$ 的 t 的可归约位置 (redex occurrence) 的集合, 即 $\mathcal{R}(t)$ 是极小集合, 满足: 对任意的 $u \in \mathcal{R}(t)$, 存在 $R_1 \cup R_2$ 中的规则 $l \rightarrow r$ 及代换 σ , 使

得 $t/u = l\sigma$ 成立;

2) $\mathcal{R}\mathcal{B}(t)$ 是满足:

$$\mathcal{R}\mathcal{B}(t) = \{u \in \mathcal{O}(t) \mid \exists P \in \mathcal{P} \exists c_1, \dots, c_n \in$$

$$\mathcal{F}(C) X_P = \{x_1, \dots, x_n\} \wedge t/u$$

$$= P\{c_1/x_1, \dots, c_n/x_n\}\}$$

的集合。

$\mathcal{R}\mathcal{B}(t)$ 的元素 u 是关于 $R_1 \cup R_2$ 的 t 的这样的可归约位置: 当重写 t/u 时, 可以把替换限制在基项的范围。我们将利用以下从定义直接得到的结论:

命题 2 对于任意的项 $t \in \mathcal{F}(G)$, $\mathcal{R}\mathcal{B}(t) \subseteq R(t)$ 。

引理 1 设 $\forall t \in \mathcal{F}(G) NF_{R_1}(t) \subseteq \mathcal{F}(C) \wedge NF_{R_2}(t) \subseteq \mathcal{F}(C)$ 。对于任意的项 $t \in \mathcal{F}(G)$, 若 $\mathcal{R}\mathcal{B}(t) = \emptyset$, 则 $t \in \mathcal{F}(C)$ 。

证明: (用反证法)。假设 $t \notin \mathcal{F}(C)$ 。由引理条件, 有 $\mathcal{R}(t) \neq \emptyset$ 。进一步, 由引理条件, 存在 $u \in \mathcal{R}(t)$, 使得

$$t/u \in \mathcal{F}(C) \wedge \forall v \in \mathcal{R}(t) (u < v \Rightarrow t/v \in \mathcal{F}(C))$$

成立。那么, 由 $\mathcal{R}\mathcal{B}(t)$ 的定义, $u \in \mathcal{R}\mathcal{B}(t)$ 。因此, $\mathcal{R}\mathcal{B}(t) \neq \emptyset$ 。由此, 引理得证。

定理 1 设 R_1 和 R_2 是满足以下条件的项重写系统:

$$1) R_1 \text{ 和 } R_2 \text{ 具有基合流性及弱基终止性;}$$

$$2) \forall t \in \mathcal{F}(C) NF_{R_1}(t) = NF_{R_2}(t);$$

$$3) \forall t \in \mathcal{F}(G) NF_{R_1}(t) \subseteq \mathcal{F}(C) \wedge NF_{R_2}(t) \subseteq \mathcal{F}(C);$$

4) 对所有的 $P \in \mathcal{P}$, 当 $X_P = \{x_1, \dots, x_n\}$ 时, 对所有的 $c_1, \dots, c_n \in \mathcal{F}(C)$, 有

$$NF_{R_1}(P\{c_1/x_1, \dots, c_n/x_n\}) \\ = NF_{R_2}(P\{c_1/x_1, \dots, c_n/x_n\})$$

成立。

$$\text{这时, } R_1 \cong R_2(\mathcal{F}(G)).$$

证明: 设 R_1 和 R_2 是满足定理条件的项重写系统。令 $\leq_D$ 是 $\mathcal{N}^2$ 上的字典序, 其中, $\mathcal{N}$ 是自然数集合。对任意的项 $t \in \mathcal{F}(G)$, 令 $B(t) = (\max_{u \in \mathcal{R}\mathcal{B}(t)} |u|, m)$, 其中, $|u|$ 是 u 的长度, m 是满足 $v = \max_{u \in \mathcal{R}\mathcal{B}(t)} |u|$ 的 $v \in \mathcal{R}\mathcal{B}(t)$ 的个数; 当 $\mathcal{R}\mathcal{B}(t) = \emptyset$ 时, 令 $B(t) = (0, 0)$ 。我们对 $B(t)$ 归纳证明 $\forall t \in \mathcal{F}(G) NF_{R_1}(t) = NF_{R_2}(t)$ 。

基底部分 若 $B(t) = (0, 0)$, 则 $\mathcal{R}\mathcal{B}(t) = \emptyset$ 。由条件 3), 我们可以应用引理 1 得出 $t \in \mathcal{F}(C)$ 。因此, 由条件 2), $NF_{R_1}(t) = NF_{R_2}(t)$ 成立。

归纳部分 设 $B(t) = (k, m)$ 且 $m \neq 0$ 。这时, 存在 $u \in \mathcal{R}\mathcal{B}(t)$, 使得 $|u| = k$ 。由 $\mathcal{R}\mathcal{B}(t)$ 的定义及条件 1) 和条件 4), 存在项 $s \in NF_{R_1}(t/u) \cap NF_{R_2}(t/u)$ 。由条件 3), $s \in \mathcal{F}(C)$ 。由此, 令 $t' = t[u \leftarrow s]$ 时, $B(t') <_D B(t)$,

且 $t \rightarrow_{R_1} t', t \rightarrow_{R_2} t'$ 。由归纳假设及条件 1), $NF_{R_1}(t') = NF_{R_2}(t') = d$ 。因此, 存在项 $s' \in NF_{R_1} \cap NF_{R_2}$ 使得 $t' \rightarrow_{R_1} s', t' \rightarrow_{R_2} s'$ 。由 $t \rightarrow_{R_1} t', t \rightarrow_{R_2} t'$ 及 R_1 和 R_2 的基合流性, 我们得到 $NF_{R_1}(t) = NF_{R_2}(t)$ 。因此, 定理得证。

注释

① 在上述定理中, 条件 1) 是关于项重写系统的性质, 当作为程序考虑项重写系统时, 条件 1) 是必要的, 它保证(对基项的)运算结果的存在性和唯一性; 同时, 它也是等价性自动证明时需要的, 因为我们在等价性自动证明时要利用项重写系统下的重写演算。

我们允许构造子项出现于重写规则的左部, 但要求满足条件 2)。条件 2) 通常是显见的, 条件 4) 是本定理的关键, 使得我们可以用归纳的手法验证它, 对于一些特殊的情况, 我们可以利用 Toyama^[7]所示的那样, 使用归纳的方法证明条件 3)。

② 为了运用定理 1, 对于任意的带有 n 个参数的 $f \in G$, $\mathcal{D}$ 必须含有型为 $f(x_1, \dots, x_n)$ 的项。但是, 从 $\mathcal{D}$ 的定义及定理 1 的证明可以看出, 我们可以对定理 1 进行一般化以处理更复杂的情况。例如, 通过对定理 1 进行一般化, 我们可以证明以下项重写系统 R_1 和 R_2 的等价性:

$$\begin{aligned} R_1 & \begin{cases} f(g(0), g(0)) & \rightarrow 0 \\ f(g(s(x)), g(s(x))) & \rightarrow s(x) \end{cases} \\ R_2 & \{f(g(x), g(x)) \rightarrow x\} \end{aligned}$$

3 证例

考虑以下计算用 $\mathcal{T}(\{0, s\})$ 表示的自然数集 $\mathcal{N}$ 上的两倍函数 $d(n) = 2 * n: \mathcal{N} \rightarrow \mathcal{N}$ 的项重写系统 R_1 和 R_2 :

$$\begin{aligned} R_1 & \begin{cases} d(0) & \rightarrow 0 \\ d(s(x)) & \rightarrow s(s(d(x))) \end{cases} \\ R_2 & \begin{cases} d(x) & \rightarrow if(x, 0, s(s(d(x-s(0)))) \\ if(0, y, z) & \rightarrow y \\ if(s(x), y, z) & \rightarrow z \\ x \rightarrow 0 & \rightarrow x \\ s(x) - s(y) & \rightarrow x - y \end{cases} \end{aligned}$$

下面利用定理 1 证明 $R_1 \cong R_2(\mathcal{T}(G))$, 其中, $C = \{0, s\}$, $G = \{0, s, d\}$ 。

因为 R_1 和 R_2 是左线性且无重叠的, 作为广为人知的结果, 它们具有合流性^[4], 因此也具有基合流性。 R_1 和 R_2 的弱终止性也是易见的。所以条件 1) 成立。因为 R_1 和 R_2 中不存在左部为构造子项的重写规则, 所以条件 2) 成立。

由定义, $\mathcal{D} = \{d(x)\}$ 。因此, 运用 Toyama^[7]的结果, 可以通过验证: 对于所有的 $c \in \mathcal{T}(C)$

$$NF_{R_1}(d(c)) = NF_{R_2}(d(c)) \subseteq \mathcal{T}(C) \quad (1)$$

证明条件 3) 和条件 4)。

为了证明 (1) 式, 对 $\mathcal{T}(C)$ 使用归纳法。

基底部分 因为 $d(0) \rightarrow_{R_1} 0, d(0) \rightarrow_{R_2} 0$, 所以 (1) 式对 $c=0$ 成立。

归纳部分 假设 (1) 式对 $c \in \mathcal{T}(C)$ 成立, 证明 (1) 式对 $s(c)$ 也成立。重写 $d(s(c))$, 我们有 $d(s(c)) \rightarrow_{R_1} s(s(d(c)))$ 及 $d(s(c)) \rightarrow_{R_2} if(s(c), 0, s(s(d(s(c)-s(0)))) \rightarrow_{R_2} s(s(d(s(c)-s(0)))) \rightarrow_{R_2} s(s(d(c)))$ 。由归纳法假设

$$NF_{R_1}(d(c)) = NF_{R_2}(d(c)) \subseteq \mathcal{T}(C)$$

所以, (1) 式对 $s(c)$ 成立。因此, 由归纳法, 条件 3) 和条件 4) 成立。

结语 运用定理 1, 我们可以利用结构归纳法及覆盖集合归纳法证明两个项重写系统在受限领域中的等价性。本方法简单有效且可以进一步推广。可以期待, 通过推广可以处理仅对 $\mathcal{D}$ 中的项具有合流性和弱终止性的项重写系统、及部分定义的函数。

参考文献

- 1 Burstall R.M, Darlington J. A transformation system for developing recursive programs. J. ACM, 1977, 24: 44~67
- 2 Feng S, et al. Dynamic term rewriting calculus and its application to inductive equational reasoning. In: Proc. DISCO '93, LNCS 722, 1993, 256~271
- 3 Feng S, et al. Mechanizing explicit inductive equational reasoning by DTRC. IEICE Trans. Inf. & Syst., 1995, E78-D(2): 113~121
- 4 Huet G. Confluent reductions: abstract properties and applications to term rewriting systems. J. ACM, 1980, 27: 797~821
- 5 Kapur D, Subramaniam M. Automating induction over mutually recursive functions. In: AMAST '96, LNCS 1101, 1996, 117~131
- 6 Knuth D E, Bendix P. Simple word problems in universal algebra. In: Leach J, Ed. Computational Problems in Abstract Algebra. Pergamon Press, Elmsford, N. Y., 1970, 263~297
- 7 Toyama Y. How to prove equivalence of term rewriting systems without induction. Theoret. Comput. Sci., 90 (1991), 1990, 369~390