《软件分析与验证》

霍尔证明系统



贺飞 清华大学软件学院

2022年3月28日

IMP 语法:

- 算术表达式、布尔表达式
- 程序语句

IMP 语义:

- 算术表达式和布尔表达式的语义
- 程序语句的语义

如何刻画对 IMP 程序的正确性需求?

——程序规约

如何证明 IMP 程序的正确性?

——霍尔证明系统

1. 程序规约

2. 霍尔证明系统

程序规约

```
while (!(y == 0)) {
    if (y >= 0) {
        y = y - 1;
    } else {
        y = y + 1;
    }
    x = x + 1;
}
```

- 考虑左边的 Pxu 程序
- 可利用前置与后置条件表达程序的 正确性规约:
 - 前置条件: $x \times y \ge 0$
 - 后置条件: x≥0
- 需要一套既能表达程序,又能刻画 规约的逻辑系统,对程序的正确性 进行推理。

- 生于 1934 年 1 月 11 日。
- 英国计算机科学家, 图灵奖得主。
- 重要贡献:
 - 快速排序算法(Quicksort)
 - 快速选择算法 (Quickselect)
 - 霍尔逻辑 (Hoare logic)
 - 通信顺序进程 (Communicating sequential process, CSP)



定义(霍尔三元组)

霍尔三元组 (hoare triple) 是形如

$$\{\varphi\}$$
 st $\{\psi\}$

的式子,其中 st 是程序, φ 和 ψ 是逻辑公式,分别称为前置条件和后置条件。

霍尔三元组的含义:从任何满足 φ 的前状态出发执行 st,如果 st 终止,那么后状态必定满足 ψ 。

霍尔三元组所刻画的程序行为中,不包括

- 不终止的程序行为
- 从不满足 φ 的状态出发的程序行为

定义(后像)

设 R 是定义在集合 X 上的一个二元关系, $Y \subseteq X$ 是 X 的一个子集,Y 关于 R 的后像(post-image)定义为:

$$post(Y,R) ::= \{x \in X \mid$$
存在 $y \in Y$ 使得 $(y,x) \in R\}$

例

考虑整数集上的小于关系 R, 即 $R = \{(a, b) \in \mathbb{Z} \times \mathbb{Z} \mid a < b\}$ 令 $Y = \{1, 2, \dots, 10\}$, 则

$$\mathit{post}(\mathit{Y}, R) = \{ y \in \mathbb{Z} \mid y \geq 2 \}$$

定义(有效的霍尔三元组)

给定程序 st 及其前置条件 φ 和后置条件 ψ , 如果

$$post(\{\varphi\}, [\![st]\!]) \subseteq \{\psi\}$$

成立,则称程序 st 满足规约 (φ,ψ) ,记作 $st\models(\varphi,\psi)$;也称霍尔三元组 $\{\varphi\}$ st $\{\psi\}$ 是有效式,记作 $\models\{\varphi\}$ st $\{\psi\}$ 。

请写出下列霍尔三元组的含义:

- 1. $\{\top\}$ st $\{\psi\}$:
- 2. $\{0 < x\}$ while (0 < x) $\{x := x + 1\}$ $\{\bot\}$:
- 3. $\{\top\}$ st $\{\bot\}$:

下面的霍尔三元组是有效式吗?

- 1. $\{x=2\}$ while (0 < x) $\{x := x-1\}$ $\{x=0\}$
- 2. $\{x = 0 \land y = 1\}$ x := x + 1 $\{x = 1 \land y = 2\}$
- 3. $\{\top\}$ while (0 < x) $\{x := x 1\}$ $\{x \le 0\}$
- 4. $\{0 < x\}$ while (0 < x) $\{x := x + 1\}$ $\{x > 0\}$

霍尔证明系统

\mathcal{S}_{PL}	\mathcal{S}_{FOL}	霍尔证明系统
推导有效命题逻	推导有效一阶逻	推导有效霍尔三
辑公式	辑公式	元组
⊢ F	⊢ <i>φ</i>	$\vdash \{\varphi\} \ st \ \{\psi\}$

推导树 17 / 44

定义(推导树)

推导树 (derivation tree) 是一颗满足下列条件的树:

- 每个节点对应一个霍尔三元组,树根对应 $\vdash \{\varphi\}$ st $\{\psi\}$,树叶对应空;
- 每个中间节点是霍尔证明系统中某条规则的结论,而它的子节 点是该规则的前提。

定义 (可推导)

如果存在以 \vdash { φ } st { ψ } 为根节点的推导树,就说霍尔三元组 { φ } st { ψ } 可推导,记作 \vdash { φ } st { ψ }。

推理规则: 空语句

18 / 44

(空语句)
$$\overline{\{\varphi\}}$$
 skip $\{\varphi\}$

引理 (空语句规则的可靠性)

霍尔三元组 $\{\varphi\}$ skip $\{\varphi\}$ 是有效式。

证明.

根据空操作语句的语义,显然成立。

(赋值)
$$\frac{}{\{\varphi[x\mapsto e]\}\ x:=e\ \{\varphi\}}$$

其中, $\varphi[x\mapsto e]$ 是指将 φ 中所有 x 的自由出现替换为 e 所得的公式。例如, $(x>0)[x\mapsto x+1]$ 的结果是 x+1>0。

为何不是下面的规则?

(赋值 ×)
$$\overline{ \{\varphi\} \ x := e \ \{\varphi[x \mapsto e]\} }$$

示例: 给定语句 x := x + 1,假设执行这条语句之前满足 x = y,根据(赋值 ×)规则,执行这条语句之后应该满足 x + 1 = y。这显然不对,正确的结果应该是 x = y + 1。

(赋值) 规则可以理解为: 如果公式 φ 在将 x 赋值为 e 后成立, 那 么公式 φ 在将 x 替换成 e 后也成立。

示例 20 / 44

以下霍尔三元组的前置条件 φ 是什么?

1.
$$\vdash \{\varphi\} \ x := x + 1 \ \{x = n\}$$

2. $\vdash \{\varphi\} \ x := 1 \ \{x = y\}$
3. $\vdash \{\varphi\} \ x := 1 \ \{y = 3\}$
4. $\vdash \{\varphi\} \ x := x^2 + 1 \ \{xy > x + y\}$

$$(x^2 + 1)y > (x^2 + 1) + y$$

引理(赋值规则的可靠性)

霍尔三元组 $\{\varphi[x\mapsto e]\}$ x:=e $\{\varphi\}$ 是有效式。

回顾赋值语句的语义

$$[\![x := e]\!] = \{(s, s') \mid s' = s[x \mapsto [\![e]\!]_s]\}$$

引理

设x为变元,e为项, φ 为公式。给定两组赋值 ρ_1, ρ_2 ,如果 $\rho_2 = \rho_1[x \mapsto [\![e]\!]_{\rho_1}]$,则 $[\![\varphi[x \mapsto e]\!]_{\rho_1} \Leftrightarrow [\![\varphi]\!]_{\rho_2}$ 。

证明.

令 $s \in \{\varphi[x \mapsto e]\}$,设 $(s, s') \in [x := e]$,即从 s 出发执行 x := e 之后得到的后状态为 s'。

根据赋值语句的语义, $s' = s[x \mapsto \llbracket e \rrbracket_s]$ 。

根据上面的引理, $\llbracket \varphi \rrbracket_{s'} = true$,即 $s' \in \{\varphi\}$ 。

我们无法证明: $\vdash \{y = 0\} \ x := 1 \ \{x = 1\}$ 但却可以证明: $\vdash \{1 = 1\} \ x := 1 \ \{x = 1\}$

注意后者的前置条件 1=1 是永真式,相当于我们在没有任何假设的情况下证明了执行 x:=1 语句后一定满足 x=1,那么当有一些假设的时候,该结论显然也成立。

前提加强规则:

(前提加强)
$$\frac{\{\varphi'\} \ st \ \{\psi\}}{\{\varphi\} \ st \ \{\psi\}} \quad \text{如果 } \varphi \models \varphi'$$

例

霍尔三元组 $\{y=0\}$ x:=1 $\{x=1\}$ 是否可推导?

类似于前提加强,还可以有结论弱化规则:

$$(结论弱化) \ \frac{\{\varphi\} \ st \ \{\psi'\}}{\{\varphi\} \ st \ \{\psi\}} \ \ \text{如果 } \psi' \models \psi$$

直观理解,如果能证明一个包含了许多事实的**更强陈述**,则必定也 能证明一个包含更少事实的**更弱陈述**。

(分支)
$$\frac{\{\varphi \land p\} \ st_1 \ \{\psi\} \quad \{\varphi \land \neg p\} \ st_2 \ \{\psi\}}{\{\varphi\} \ \text{if } (p) \ \{st_1\} \ \text{else} \ \{st_2\} \ \{\psi\}}$$

- 如果分支条件 p 成立,执行 st_1 分支,执行结束后要求满足后置条件 ψ 。
- 如果分支条件 p 不成立,执行 st_2 分支,执行结束后也要满足后置条件 ψ 。

例

$$\begin{split} \psi &= (x \geq 0 \to y = x - 1) \land (x < 0 \to y = x + 1) \\ \psi_1 &= (x \geq 0 \to y - 1 = x - 1) \land (x < 0 \to y - 1 = x + 1) \\ \psi_2 &= (x \geq 0 \to y + 1 = x - 1) \land (x < 0 \to y + 1 = x + 1) \end{split}$$

注意: 在应用前提加强和结论弱化到上述证明的过程中, 需要判断:

$$x = y \land y \ge 0 \models \psi_1$$

$$x = y \land y < 0 \models \psi_2$$

引理

如果霍尔三元组 $\{\varphi \land p\}$ st_1 $\{\psi\}$ 和 $\{\varphi \land \neg p\}$ st_2 $\{\psi\}$ 都是有效式,则霍尔三元组 $\{\varphi\}$ **if** (p) $\{st_1\}$ **else** $\{st_2\}$ $\{\psi\}$ 是有效式。

回顾分支语句的语义

$$\llbracket \mathbf{if} (p)\{st_1\} \mathbf{else} \{st_2\} \rrbracket = \left\{ (s, s') \middle| \begin{array}{c} \llbracket p \rrbracket_s = true \ \amalg \ (s, s') \in \llbracket st_1 \rrbracket \\ \stackrel{\mathbf{I}}{\text{!!}} \llbracket p \rrbracket_s = false \ \coprod \ (s, s') \in \llbracket st_2 \rrbracket \right\}$$

证明.

令 $s \in \{\varphi\}$, 设 $(s, s') \in [[if(p) \{st_1\} else \{st_2\}]]$, 需要证明 $s' \in \{\psi\}$ 。 分两种情况讨论:

- 如果 $[p]_s = true$ 且 $(s, s') \in [st_1]$,根据 $\{\varphi \land p\}$ st_1 $\{\psi\}$ 是有效式知 $s' \in \{\psi\}$ 。
- 如果 $[p]_s = false$ 且 $(s, s'') \in [st_2]$,根据 $\{\varphi \land \neg p\}$ st_2 $\{\psi\}$ 是有效式知 $s'' \in \{\psi\}$ 。

(順序)
$$\frac{\{\varphi_1\} \ st_1 \ \{\varphi_2\} \ \ \{\varphi_2\} \ \ st_2 \ \{\varphi_3\}}{\{\varphi_1\} \ \ st_1; st_2 \ \{\varphi_3\}}$$

- 注意: 需要找到合适的逻辑公式 φ_2 使两个前提都成立。
- 在这个过程中, 很可能需要用到前提加强和结论弱化规则。

$$\begin{cases} x = x' \\ \land y = y' \end{cases} t := x \begin{cases} t = x' \\ \land y = y' \end{cases} x := y \begin{cases} t = x' \\ \land x = y' \end{cases}$$

$$\begin{cases} x = x' \\ \land y = y' \end{cases} t := x; x := y \begin{cases} t = x' \\ \land x = y' \end{cases}$$

$$\begin{cases} t = x' \\ \land x = y' \end{cases} t := x; x := y; y := t \begin{cases} y = x' \\ \land x = y' \end{cases}$$

$$\begin{cases} x = x' \\ \land x = y' \end{cases} t := x; x := y; y := t \begin{cases} y = x' \\ \land x = y' \end{cases}$$

引理(顺序规则的可靠性)

如果霍尔三元组 $\{\varphi_1\}$ st_1 $\{\varphi_2\}$ 和 $\{\varphi_2\}$ st_2 $\{\varphi_3\}$ 都是有效式,则霍尔三元组 $\{\varphi_1\}$ $st_1; st_2$ $\{\varphi_3\}$ 是有效式。

回顾顺序语句的语义

$$[st_1; st_2] = \{(s, s') \mid$$
存在 s'' 使得 $(s, s'') \in [st_1], (s'', s') \in [st_2]\}$

证明.

令 $s \in \{\varphi_1\}$,设 $(s,s') \in [st_1; st_2]$,需要证明 $s' \in \{\varphi_3\}$ 。 根据顺序语句的语义,存在 s'' 使得 $(s,s'') \in [st_1]$ 及 $(s'',s') \in [st_2]$ 。 根据 $\{\varphi_1\}$ st_1 $\{\varphi_2\}$ 和 $\{\varphi_2\}$ st_2 $\{\varphi_3\}$ 都是有效式,有 $s'' \in \{\varphi_2\}$, $s' \in \{\varphi_3\}$ 。 到此为止,我们已经分别针对空语句、赋值语句、分支语句和顺序语句引入了推理规则。应用这些推理规则会让待证霍尔三元组中的程序**至少减少一个构造子**,从而起到简化的作用。

对于循环,容易得到下面的结论:

引理

语句 while $(p)\{st\}$ 和语句 if $(p)\{st;$ while $(p)\{st\}\}$ else skip 是语义 等价的。

继而可得下面的推理规则:

(循环展开)
$$\frac{\{\varphi \land p\} \ \ \textbf{if} \ (p)\{st; \textbf{while} \ (p)\{st\}\} \ \textbf{else skip} \ \ \{\psi\}}{\{\varphi\} \ \ \textbf{while} \ (p)\{st\} \ \ \{\psi\}}$$

该推理规则的本质是将循环展开,但展开后的程序仍然包含原来的 while 循环。包含这条规则的证明系统将无法保证整个证明过程 (即构建推导树的过程) 的终止性。

(循环)
$$\frac{\{\varphi \wedge p\} \ st \ \{\varphi\}}{\{\varphi\} \ \mathbf{while} \ (p)\{st\} \ \{\varphi \wedge \neg p\}}$$

- φ 是一个循环不变式:
 - 首次到达循环头位置, φ 成立;
 - 从满足 φ 的状态出发,进入循环并迭代一次之后, φ 仍然成立。
- 该规则刻画了循环不变式的第二个条件,第一个条件(即程序执行到循环语句时, φ 成立)需要与其他规则配合来检查。

任务:证明 $\{x \ge 42 \land y \le -23\}$ P_{xy} $\{x \ge 53\}$ 是有效霍尔三元组。

以 if-st 代表: if $(y \ge 0)$ {y := y - 1} else {y := y + 1} 以 P_{xy} 代表: while $(y \ne 0)$ {if-st; x := x + 1}

順序
$$\frac{\{\psi_1 \wedge (y \neq 0)\} \text{ if-st } \{\psi_2\}}{\{\psi_1 \wedge (y \neq 0)\} \text{ if-st;} x := x + 1 \ \{\psi_1\}}$$
加强
$$\frac{\{\psi_1 \wedge (y \neq 0)\} \text{ if-st;} x := x + 1 \ \{\psi_1\}}{\{\psi_1 \wedge \neg (y \neq 0)\}}$$
弱化
$$\frac{\{x \geq 42 \wedge y \leq -23\} P_{xy} \{\psi_1 \wedge \neg (y \neq 0)\}}{\{x \geq 42 \wedge y \leq -23\} P_{xy} \{x \geq 53\}}$$

其中:

$$\psi_1 = x - y \ge 53$$
 (循环不变式)
 $\psi_2 = x + 1 - y \ge 53$

其中:

$$\psi_1 = x - y \ge 53$$
 (循环不变式)
 $\psi_2 = x + 1 - y \ge 53$
 $\psi_2' = x + 1 - (y - 1) \ge 53$
 $\psi_2'' = x + 1 - (y + 1) \ge 53$

规则可靠性 37 / 44

引理(循环规则的可靠性)

如果霍尔三元组 $\{\varphi \land p\}$ st $\{\psi\}$ 是有效式,则霍尔三元组 $\{\varphi\}$ while $\{p\}$ $\{st\}$ $\{\psi\}$ 是有效式。

至此,我们证明了霍尔逻辑系统中每一条推理规则的可靠性。

定理(霍尔证明系统的可靠性)

霍尔证明系统是可靠的 (sound), 即通过该证明系统推导出的所有霍尔三元组都是有效式。换句话说,如果存在一颗以 $\{\varphi\}$ st $\{\psi\}$ 为根节点的推导树(记作 \vdash $\{\varphi\}$ st $\{\psi\}$),则该霍尔三元组必是有效式(记作 \models $\{\varphi\}$ st $\{\psi\}$),此时也称程序 st 满足规约 (φ,ψ) 。

证明:设推导树的高度为 n。对推导树的高度进行归纳以证明下面的结论:推导树根节点对应的霍尔三元组一定是有效式。

• 当 n=1 时,设根节点被标注 $\{\varphi\}$ st $\{\psi\}$,其他节点都是叶子节点,且标注为空。根据推导树的定义,从根节点到叶子节点一定应用了前提为空的推理规则,只能是空语句规则或者赋值规则。根据这两条规则的可靠性可知结论成立。

• 假设结论在 $n \le k$ 时成立,需要证明在 n = k+1 时也成立。 设推导树的根节点被标注 $\{\varphi\}$ st $\{\psi\}$,其子节点分别被标注 $\{\varphi_1\}$ st₁ $\{\psi_1\}$, ..., $\{\varphi_m\}$ st_m $\{\psi_m\}$ 。根据假设,所有子节点都 是高度 $\le k$ 的推导树的根节点,所以这些子节点上标注的霍尔 三元组都是有效式。根据推导树的定义,

$$\{\varphi_1\} \ st_1 \ \{\psi_1\} \ \dots \ \{\varphi_m\} \ st_m \ \{\psi_m\}$$

$$\{\varphi\} \ st \ \{\psi\}$$

必须是某条推理规则的实例。能够在此处应用的规则只能是前提加强、结论弱化、顺序、分支和循环规则中的一条。根据这些规则的可靠性,可证根节点的霍尔三元组 $\{\varphi\}$ st $\{\psi\}$ 一定是有效式。

定理(霍尔证明系统的完备性)

霍尔证明系统是相对完备的 (relatively complete): 如果存在一个一阶逻辑判定工具,则通过该系统可以推导出 IMP 语言中所有有效的霍尔三元组。

- 回忆一下前面学到的前提加强和结论弱化规则,我们需要证明 形如 $p \to q$ 的一阶逻辑蕴涵式的有效性。
- 对于一个有效的霍尔三元组,能否用霍尔证明系统推导依赖于证明过程中所产生的一阶逻辑蕴涵式能否得到有效判定。
- 不幸的是,IMP 语言中允许乘法,对应的 T_{PA} 算术理论是不可判定的!
- 对于更复杂的语言,相对完备性可能也无法保证。

基本思路是减少证明过程的不确定性:

- 为每个循环选择一个"合适"的循环不变式
- 结论弱化规则仅用于等价变换(如公式的语法变形)
- 从后往前应用顺序组合规则
- 仅对循环之前的语句应用前提严格加强规则
- 如果选择的循环不变式不足以完成证明,则选择新的循环不变式,并重新开始推导。

应用的难点和关键是选择合适的循环不变式。 然而,生成合适的循环不变式的问题是不可判定问题(否则,程序 验证将可判定!)。 总结: 霍尔证明系统

42 / 44

IMP 程序规约:

- 前置条件、后置条件、后像
- 霍尔三元组

IMP 霍尔证明系统:

- 推理规则
- 可靠、相对完备

- 进一步理解循环
- 包含数据结构的程序的验证

谢谢!