بهبود روش وارسی مدل با استفاده از تعبیر مجرد

پویا پرتو

دکتر مجید علی‌زاده و دکتر مجتبی محتهدی

دانشکده‌ی ریاضی، آمار و علومِ کامپیوتر
دانشگاهِ تهران

شهریور ۱۴۰۲

پیشگفتار

درستی‌یابی برنامه‌های کامپیوتری

- اهمیت درستی‌یابی برنامه‌ها
- روش‌های پویا
- روش‌های ایستا

روش‌های ایستا

- روش‌های صوری: استفاده از نظریه‌های مختلف در منطق ریاضی
- قضیه‌ی رایس: تصمیم ناپذیری مسئله در حالت کلی
- روش وارسی مدل: استفاده از منطق‌های وجهی مختلف در کنار مدل کردن برنامه با کمک معناشناسی منطق
- روش‌های استنتاجی: استفاده از نظریه نوع یا منطق تفکیک (منطق هور) و اثبات‌یارها (مانند Coq, Isabelle و ...)
- استفاده از نظریه تعبیر مجرد: تخمین زدن **درست** معناشناسی برنامه‌های کامپیوتری

هدف

- بیان دوباره روش وارسی مدل با استفاده از عبارات منظم به جای منطق‌های وجهی
- فایده: آشنا بودن برنامه نویسان با عبارات منظم
- روش وارسی مدل در سه صورت جدید بیان می‌شود.

مفاهیم اولیه

منطق زمانی خطی

نحو

$$\phi \in \Phi \Leftrightarrow \phi ::= \pi | \phi \vee \phi | \neg\phi | \bigcirc \phi | \phi \mathcal{U} \phi \qquad (\pi \in \Pi)$$

معناشناسی

$$M, i \models \pi \quad iff \quad \pi \in M(i),$$

$$M, i \models \neg\phi \quad iff \quad M, i \not\models \phi,$$

$$M, i \models \phi \vee \psi \quad iff \quad M, i \models \phi \quad or \quad M, i \models \psi,$$

$$M, i \models \bigcirc\phi \quad iff \quad M, i+1 \models \phi,$$

$$M, i \models \phi \mathcal{U} \psi \quad iff \quad \exists k \geq i \in \mathbb{N}_0 : \forall i \leq j < k : M, j \models \phi \quad and \quad M, k \models \psi.$$

نحو زبان مورد بررسی

$$
\begin{aligned}
&x, y, \ldots \in \mathbb{X},\\
&A \in \mathbb{A} ::= 1 \mid x \mid A_1 - A_2\\
&B \in \mathbb{B} ::= A_1 < A_2 \mid B_1 \text{ nand } B_2\\
&E \in \mathbb{E} ::= A \mid B\\
&S \in \mathbb{S} ::=\\
&\qquad x \doteq A;\\
&\qquad \mid \ ;\\
&\qquad \mid \ \text{if } (B)\ S \mid \text{if } (B)\ S \text{ else } S\\
&\qquad \mid \ \text{while } (B)\ S \mid \text{break};\\
&\qquad \mid \ \{Sl\}\\
&Sl \in \mathbb{Sl} ::= Sl \ \ S \mid \ \text{ə}\\
&P \in \mathbb{P} \ ::= Sl
\end{aligned}
$$

برچسب‌ها

$[\![\text{at}]\!][\![S]\!]$: برچسب شروع S.

$[\![\text{aft}]\!][\![S]\!]$: برچسب پایان S، اگر پایانی داشته باشد.

$[\![\text{esc}]\!][\![S]\!]$: یک مقدار بولی را بازمی‌گرداند که بسته به اینکه در S عبارت دستوری break; وجود دارد یا خیر، مقدار درست یا غلط را برمی‌گرداند.

$[\![\text{brk-to}]\!][\![S]\!]$: برچسبی است که اگر حین اجرای S عبارت دستوری break; اجرا شود، برنامه از آن برچسب ادامه پیدا می کند.

$[\![\text{brks-of}]\!][\![S]\!]$: مجموعه‌ای از برچسب‌های عبارت‌های دستوری break; که داخل S هستند را برمی‌گرداند.

$[\![\text{in}]\!][\![S]\!]$: مجموعه‌ای از تمام برچسب‌های درون S را برمی‌گرداند.

$[\![\text{labs}]\!][\![S]\!]$: مجموعه‌ای از تمام برچسب‌هایی که با اجرای S قابل دسترسی هستند را برمی‌گرداند.

مجموعه‌ی همه‌ی برچسب‌ها را با $\mathbb{L}$ نشان می‌دهیم.

معناشناسی برنامه‌ها

محیط

به ازای مجموعه مقادیر $\mathbb{V}$ و مجموعه متغیرها $\mathbb{X}$ تابع $\rho : \mathbb{X} \to \mathbb{V}$ را یک محیط می‌گوییم. مجموعه‌ی همه‌ی محیط‌ها را با $\mathbb{EV}$ نمایش می‌دهیم.

وضعیت

به ازای مجموعه مقادیر(وضعیت): به هر زوج مرتب متشکل از یک برچسب l و یک محیط ρ یک وضعیت $\langle l, \rho \rangle$ می‌گوییم. مجموعه‌ی همه‌ی وضعیت‌ها را با $\mathfrak{S}$ نشان می‌دهیم.

رد پیشوندی

به یک دنباله از وضعیت‌ها(با امکان تهی بودن) یک رد پیشوندی می‌گوییم.

عملگر چسباندن

برای $\pi_1, \pi_2 \in \mathfrak{S}^{+\infty}$ و $\sigma_1, \sigma_2 \in \mathfrak{S}$ داریم:

$\blacktriangleleft\ \pi_1 \in \mathfrak{S}^{\infty}$:

$$\pi_1 \bowtie \pi_2 = \pi_1$$

$\blacktriangleleft\ \pi_1 \in \mathfrak{S}^{+}$:

$\blacktriangleleft\blacktriangleleft\ \sigma_1 = \sigma_2$:

$$\pi_1\sigma_1 \bowtie \sigma_2\pi_2 = \pi_1\sigma_1\pi_2$$

$\blacktriangleleft\blacktriangleleft\ \sigma_1 \neq \sigma_2$:

در این حالت $\pi_1 \bowtie \pi_2$ تعریف نمی‌شود.

معنای عبارت‌های حسابی - تابع $\mathcal{A}$

تابع $\mathcal{A} : \mathbb{A} \to (\mathbb{EV} \to \mathbb{V})$ را به صورت بازگشتی روی ساختار $\mathrm{A} \in \mathbb{A}$ به شکل زیر تعریف می‌کنیم:

$$\blacktriangleleft \ \mathcal{A}[\![1]\!]\rho = 1$$

$$\blacktriangleleft \ \mathcal{A}[\![\mathrm{x}]\!]\rho = \rho(\mathrm{x})$$

$$\blacktriangleleft \ \mathcal{A}[\![\mathrm{A}_1 - \mathrm{A}_2]\!]\rho = \mathcal{A}[\![\mathrm{A}_1]\!]\rho - \mathcal{A}[\![\mathrm{A}_2]\!]\rho$$

معنای عبارت‌های بولی - تابع $\mathcal{B}$

تابع $\mathcal{B} : \mathbb{B} \to (\mathbb{EV} \to \mathbb{BOOL})$ را به صورت بازگشتی روی ساختار $\mathrm{B} \in \mathbb{B}$ به شکل زیر تعریف می‌کنیم:

- اگر $\mathcal{A}[\![\mathrm{A}_1]\!]\rho$ کوچکتر از $\mathcal{A}[\![\mathrm{A}_2]\!]\rho$ باشد $\mathcal{B}[\![\mathrm{A}_1 < \mathrm{A}_2]\!]\rho = True$
- اگر $\mathcal{A}[\![\mathrm{A}_1]\!]\rho$ بزرگتر از $\mathcal{A}[\![\mathrm{A}_2]\!]\rho$ باشد $\mathcal{B}[\![\mathrm{A}_1 < \mathrm{A}_2]\!]\rho = False$
- $\mathcal{B}[\![\mathrm{B}_1 \mathrm{nand} \mathrm{B}_2]\!]\rho = \neg(\mathcal{B}[\![\mathrm{B}_1]\!]\rho \wedge \mathcal{B}[\![\mathrm{B}_2]\!]\rho)$

معنای برنامه‌ها - تابع $\mathcal{S}^*$ (دستور مقداردهی)

$$\blacktriangleleft \mathrm{S} = \mathrm{x} \doteq \mathrm{A}; \ :$$

$$\mathcal{S}^*[\![\mathrm{S}]\!] = \{\langle at[\![\mathrm{S}]\!], \rho\rangle | \rho \in \mathbb{EV}\} \cup$$

$$\{\langle at[\![\mathrm{S}]\!], \rho\rangle \langle aft[\![\mathrm{S}]\!], \rho[\mathrm{x} \leftarrow \mathcal{A}[\![\mathrm{A}]\!]\rho]\rangle | \rho \in \mathbb{EV}\}$$

معنای برنامه‌ها - تابع $\mathcal{S}^*$ (دستور شرط)

$$\blacktriangleleft \ \mathrm{S} = \text{if } (\mathrm{B}) \ \mathrm{S_t} \text{ else } \mathrm{S_f} :$$

$$\mathcal{S}^*[\![\mathrm{S}]\!] = \{\langle at[\![\mathrm{S}]\!], \rho\rangle | \rho \in \mathbb{EV}\}$$

$$\cup\{\langle at[\![\mathrm{S}]\!], \rho\rangle\langle at[\![\mathrm{S_f}]\!], \rho\rangle\pi | \mathcal{B}[\![\mathrm{B}]\!]\rho = \mathit{False} \wedge \langle at[\![\mathrm{S_f}]\!], \rho\rangle\pi \in \mathcal{S}[\![\mathrm{S_f}]\!]\}$$

$$\cup\{\langle at[\![\mathrm{S}]\!], \rho\rangle\langle at[\![\mathrm{S_t}]\!], \rho\rangle\pi | \mathcal{B}[\![\mathrm{B}]\!]\rho = \mathit{True} \wedge \langle at[\![\mathrm{S_t}]\!], \rho\rangle\pi \in \mathcal{S}[\![\mathrm{S_t}]\!]\}$$

معنای برنامه‌ها - تابع $\mathcal{S}^*$ (دستور حلقه)

$$\mathcal{S}^*[\![\mathrm{S}]\!] = lfp^{\subseteq}\, \mathcal{F}[\![\mathrm{S}]\!],$$

$$\mathcal{F}[\![\mathrm{S}]\!]X = \{\langle at[\![\mathrm{S}]\!], \rho\rangle | \rho \in \mathbb{EV}\} \cup$$

$$\{\pi_2\langle l, \rho\rangle \langle aft[\![\mathrm{S}]\!], \rho\rangle | \pi_2\langle l, \rho\rangle \in X \wedge \mathcal{B}[\![\mathrm{B}]\!]\rho = False \wedge l = at[\![\mathrm{S}]\!]\} \cup$$

$$\{\pi_2\langle l, \rho\rangle \langle at[\![\mathrm{S_b}]\!], \rho\rangle \pi_3 | \pi_2\langle l, \rho\rangle \in X \wedge \mathcal{B}[\![\mathrm{B}]\!]\rho = True \wedge$$

$$\langle at[\![\mathrm{S_b}]\!], \rho\rangle \pi_3 \in \mathcal{S}[\![\mathrm{S_b}]\!] \wedge l = at[\![\mathrm{S}]\!]\}$$

صوری‌سازی جدید برای روش وارسی مدل

نحو عبارات منظم

مجموعه‌ی $\mathbb{R}$ توسط گرامر زیر ساخته می‌شود.

$$
\begin{aligned}
R ::= \quad & \varepsilon \\
| \quad & L : B \\
| \quad & R_1 R_2 \quad (or\ R_1 \bullet R_2) \\
| \quad & R_1 \mid R_2 \quad (or\ R_1 + R_2) \\
| \quad & R_1^* \\
| \quad & R_1^+ \\
| \quad & (R_1)
\end{aligned}
$$

معناشناسی عبارات منظم

$$\blacktriangleleft\ \mathcal{S}^r[\![\varepsilon]\!] = \{\langle \underline{\rho}, \epsilon\rangle | \underline{\rho} \in \underline{\mathbb{EV}}\}$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{L} : \mathrm{B}]\!] = \{\langle \underline{\rho}, \langle l, \rho\rangle\rangle | l \in \mathrm{L} \wedge \mathcal{B}[\![\mathrm{B}]\!]\underline{\rho}, \rho\}$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}_1\mathrm{R}_2]\!] = \mathcal{S}^r[\![\mathrm{R}_1]\!] \bowtie \mathcal{S}^r[\![\mathrm{R}_2]\!]$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}_1 \mid \mathrm{R}_2]\!] = \mathcal{S}^r[\![R_1]\!] \cup \mathcal{S}^r[\![R_2]\!]$$

معناشناسی عبارات منظم (ادامه)

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}]\!]^0 = \mathcal{S}^r[\![\varepsilon]\!],$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}]\!]^{n+1} = \mathcal{S}^r[\![\mathrm{R}]\!]^n \bowtie \mathcal{S}^r[\![\mathrm{R}]\!].$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}^*]\!] = \bigcup_{n\in\mathbb{N}} \mathcal{S}^r[\![\mathrm{R}^n]\!],$$

$$\blacktriangleleft\ \mathcal{S}^r[\![\mathrm{R}^+]\!] = \bigcup_{n\in\mathbb{N}\setminus\{0\}} \mathcal{S}^r[\![\mathrm{R}^n]\!].$$

$$\blacktriangleleft\ \mathcal{S}^r[\![(\mathrm{R})]\!] = \mathcal{S}^r[\![\mathrm{R}]\!].$$

عبارات منظم

ارضا پذیری

می‌گوییم، در محیط اولیه‌ی $\underline{\rho}$ رد پیشوندی π عبارت منظم R را ارضا می‌کند، اگر و تنها اگر $\langle \underline{\rho}, \pi \rangle \in \mathcal{S}^r[\![\mathrm{R}]\!]$.

گونه‌های مختلف عبارات منظم

عبارت منظم تهی - $\mathbb{R}_\varepsilon$

$\mathbb{R}_\varepsilon$ توسط گرامر زیر تولید می‌شود.

$$R \quad ::= \quad \varepsilon \mid R_1 R_2 \mid R_1 + R_2 \mid R_1^* \mid R_1^+ \mid (R_1)$$

گونه‌های مختلف عبارات منظم

عبارات منظم ناتهی - $\mathbb{R}^+$

$\mathbb{R}^+$ توسط گرامر زیر تولید می‌شود. توسط گرامر زیر تولید می‌شود.

$$R \quad ::= \quad L : B \mid \varepsilon R_2 \mid R_1 \varepsilon \mid R_1 R_2 \mid R_1 + R_2 \mid R_1^+ \mid (R_1)$$

گونه‌های مختلف عبارات منظم

عبارات منظم بدون انتخاب - $\mathbb{R}^\dagger$

$\mathbb{R}^\dagger$ توسط گرامر زیر تولید می‌شود.

$$R \quad ::= \quad \varepsilon \mid L : B \mid R_1 R_2 \mid {R_1}^* \mid {R_1}^+ \mid (R_1)$$

صورت جدید مسئله‌ی وارسی مدل

بستار پیشوندی

اگر $\Pi \in P(\underline{\mathbb{EV}} \times \mathfrak{S}^+)$، آنگاه بستار پیشوندی Π را به صورت زیر تعریف می‌کنیم:

$$\text{prefix}(\Pi) = \{\langle \underline{\rho}, \pi \rangle | \pi \in \mathfrak{S}^+ \wedge \exists\, \pi' \in \mathfrak{S}^* : \langle \underline{\rho}, \pi\pi' \rangle \in \Pi\}$$

صورت جدید مسئله‌ی وارسی مدل

وارسی مدل

اگر $\underline{\rho} \in \underline{\mathbb{EV}}$,$\mathrm{P} \in \mathbb{P}, \mathrm{R} \in \mathbb{R}^{+}$ آنگاه:

$$\mathrm{P}, \underline{\rho} \models \mathrm{R} \Leftrightarrow (\{\underline{\rho}\} \times \mathcal{S}^{*}[\![\mathrm{P}]\!]) \subseteq \mathrm{prefix}(\mathcal{S}^{r}[\![\mathrm{R} \bullet (? : T)^{*}]\!])$$

توقف پذیری

برنامه‌ی P را به همراه محیط اولیه $\underline{\rho}$ توقف پذیر می‌گوییم، اگر و تنها اگر وجود داشته باشد $\pi \in \mathcal{S}^*[\![\mathrm{P}]\!]$، که ($\rho$ محیط متناظر با محیط اولیه‌ی $\underline{\rho}$ است):

$$\pi = \langle at[\![\mathrm{P}]\!], \rho \rangle \pi'$$

و اینکه $\langle aft[\![\mathrm{P}]\!], \rho' \rangle$ در π حضور داشته باشد. در این صورت می‌نویسیم $\mathrm{P}, \underline{\rho} \downarrow$.

در مورد توقف پذیری

قضیه

برای برنامه‌ی P و محیط اولیه‌ی $\underline{\rho}$ داریم $\mathrm{P}, \underline{\rho} \downarrow$، اگر و تنها اگر ρ محیط متناظر با محیط اولیه‌ی $\underline{\rho}$ باشد و

$$\forall \pi \in \mathfrak{S}^{+} : \langle at[\![\mathrm{P}]\!], \rho \rangle \pi \in \mathcal{S}^{*}[\![\mathrm{P}]\!] \rightarrow \langle at[\![\mathrm{P}]\!], \rho \rangle \pi \in \mathbb{R}^{+}.$$

در مورد توقف پذیری

- داریم:

$$\mathrm{P}, \underline{\rho} \models \varepsilon$$

$$\Leftrightarrow$$

$$(\{\underline{\rho}\} \times \mathcal{S}^*[\![\mathrm{P}]\!]) \subseteq \mathrm{prefix}(\mathcal{S}^r[\![\varepsilon \bullet (? : T)^*]\!]) = \mathrm{prefix}(\mathcal{S}^r[\![(? : T)^*]\!])$$

در مورد توقف پذیری

- پس اگر الگوریتمی برای تشخیص $\mathrm{P}, \underline{\rho} \models \varepsilon$ داشته باشیم، مسئله‌ی توقف حل می‌شود.

- پس پیاده‌سازی این روش غیر ممکن است.

- دو صورت دیگر هم قابل پیاده‌سازی نیستند!

وارسی مدل منظم

وارسی مدل منظم

- ساختار عبارات منظم به صورت اضافه می‌شود.