多段階 let 挿入を行うコード生成言語の 型システムの設計

大石純平 亀山幸義

筑波大学 コンピュータ・サイエンス専攻

2016/9/9 日本ソフトウェア科学会第 33 回大会

アウトライン

- 1 概要
- 2 研究の目的
- 3 研究の内容

アウトライン

- 1 概要
- ② 研究の目的
- 3 研究の内容

概要

プログラムを生成するプログラミング言語 (=<mark>コード生成言語</mark>) の安全性を保証する研究

概要

プログラムを生成するプログラミング言語 (=<mark>コード生成言語</mark>) の安全性を保証する研究

- 効率的なコードの生成
- 安全性の保証

概要

プログラムを生成するプログラミング言語 (=コード生成言語) の安全性を保証する研究

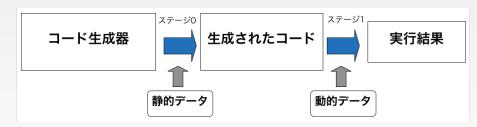
- 効率的なコードの生成
- 安全性の保証
- ⇒ <mark>多段階 let 挿入</mark>を効率的かつ安全に扱うための型システムを 構築

コード生成



- コード生成ステージとコード実行ステージ
- コード生成をサポートするプログラム言語 ⇒ コード生成言語

コード生成



- コード生成ステージとコード実行ステージ
- コード生成をサポートするプログラム言語 ⇒ コード生成言語
- 生成するプログラムだけでなく、生成されたプログラムも型の整合性が静的に(生成前に)保証される

コード生成器 🛶* 生成されるコード

コード生成器 →* 生成されるコード (<u>int</u> 3)

コード生成器
$$\leadsto^*$$
 生成されるコード
$$(\underline{\mathsf{int}}\ 3) \leadsto^* < 3 > \\ (\underline{\mathsf{int}}\ 3) + (\underline{\mathsf{int}}\ 5) \leadsto^* < 3 + 5 >$$

コード生成器
$$\leadsto^*$$
 生成されるコード (int 3) \leadsto^* <3> (int 3) $+$ (int 5) \leadsto^* <3 + 5> $\underline{\lambda}x$. x + (int 3)

コード生成器
$$\leadsto^*$$
 生成されるコード (int 3) \leadsto^* <3> (int 3) \pm (int 5) \leadsto^* <3 + 5> $\underline{\lambda}x. \ x + (int 3) \leadsto^* <\lambda x'.x' + 3>$

コード生成器
$$\leadsto^*$$
 生成されるコード
$$\underbrace{(\text{int }3) \, \leadsto^* < 3>}$$

$$\underbrace{(\text{int }3) \, \pm \, (\text{int }5) \, \leadsto^* < 3+5>}$$

$$\underline{\lambda}x. \, x \, \pm \, (\text{int }3) \, \leadsto^* < \lambda x'.x'+3>$$
 for $x=\cdots$ to \cdots do $\cdots \leadsto^*$ x'=\cdots to \cdots do $\cdots \leadsto$

コード生成器
$$\leadsto^*$$
 生成されるコード
$$(\underline{\mathsf{int}}\ 3) \leadsto^* < 3 > \\ (\underline{\mathsf{int}}\ 3) \, \pm \, (\underline{\mathsf{int}}\ 5) \leadsto^* < 3 + 5 > \\ \underline{\lambda}x.\ x \, \pm \, (\underline{\mathsf{int}}\ 3) \leadsto^* < \lambda x'.x' + 3 > \\ \mathsf{for}\ x = \cdots \ \mathsf{to} \cdots \ \mathsf{do} \ \cdots \leadsto^* < \mathsf{for}\ x' = \cdots \ \mathsf{to} \cdots \ \mathsf{do} \ \cdots >$$

゙コードコンビネータ

- 下線つきの演算子
- コードを引数にとり、コードを返す

普通の power 関数

```
\begin{aligned} \mathsf{power} &= \ \lambda x. \mathsf{fix} \ \lambda f. \lambda n. \\ &\quad \mathsf{if} \ n = 0 \ \mathsf{then} \ 1 \\ &\quad \mathsf{else} \ x \ \times \ (f \ (n-1)) \end{aligned}
```

gen_power: power コード生成器 ${\rm gen_power} = \ \underline{\lambda}x.{\rm fix} \ \lambda f.\lambda n.$ ${\rm if} \ n=0 \ {\rm then} \ (\underline{{\rm int}} \ 1)$ ${\rm else} \ x\times (f\ (n-1))$

gen_power: power **コード生成器**

$$\begin{split} \mathsf{gen_power} &= \ \underline{\lambda} x. \mathsf{fix} \ \lambda f. \lambda n. \\ &\quad \mathsf{if} \ n = 0 \ \mathsf{then} \ (\underline{\mathsf{int}} \ 1) \\ &\quad \mathsf{else} \ x \times (f \ (n-1)) \end{split}$$

n=5 に特化したコード生成:

gen_power: power コード生成器

$$\begin{split} \mathsf{gen_power} &= \ \underline{\lambda} x. \mathsf{fix} \ \lambda f. \lambda n. \\ &\quad \mathsf{if} \ n = 0 \ \mathsf{then} \ (\underline{\mathsf{int}} \ 1) \\ &\quad \mathsf{else} \ x \times (f \ (n-1)) \end{split}$$

n=5 に特化したコード生成:

gen_power: power コード生成器

$$\begin{split} \mathsf{gen_power} &= \ \underline{\lambda}x.\mathsf{fix} \ \lambda f.\lambda n. \\ &\quad \mathsf{if} \ n = 0 \ \mathsf{then} \ (\underline{\mathsf{int}} \ 1) \\ &\quad \mathsf{else} \ x \ \underline{\times} \ (f \ (n-1)) \end{split}$$

n=5 に特化したコード生成:

$$\underline{\lambda}x$$
. gen_power $x \mathrel{5} \leadsto^* \lessdot \lambda x'$. $x' \times x' \times x' \times x' \times x' \times 1 \gt$

gen_power 関数によって生成されたコードは power 関数より高速

二重 for ループのコード生成器

コード生成器

$$\begin{array}{l} \underline{\textbf{for}} \ x = (\underline{\textbf{int}} \ 0) \ \underline{\textbf{to}} \ (\underline{\textbf{int}} \ n) \ \underline{\textbf{do}} \\ \underline{\textbf{for}} \ y = (\underline{\textbf{int}} \ 0) \ \underline{\textbf{to}} \ (\underline{\textbf{int}} \ m) \ \underline{\textbf{do}} \\ \underline{\textbf{set}} \ a \ (x,y) \ \text{complex calculation} \end{array}$$

二重 for ループのコード生成器

コード生成器

生成されるコード

< for
$$x' = 0$$
 to n do
for $y' = 0$ to m do
 $a[x', y'] \leftarrow$ complex calculation >

多段階 let 挿入の例

生成されるコード

< for
$$x'=0$$
 to n do
$$a[x',y'] \leftarrow \text{complex calculation} \\ b[x',y'] \leftarrow \text{complex calculation} >$$

多段階 let 挿入の例

生成されるコード

```
< let w = \text{complex calculation in} for x' = 0 to n do let u = \text{complex calculation in} for y' = 0 to m do a[x', y'] \leftarrow u b[x', y'] \leftarrow w
```

多段階 let 挿入の例

生成されるコード

```
< let w = complex calculation in for x' = 0 to n do let u = complex calculation in for y' = 0 to m do a[x', y'] \leftarrow u b[x', y'] \leftarrow w
```

多段階 let 挿入

- 入れ子になった for ループなどを飛び越えた複数のコード移動を許す仕組み
- ループ不変式の移動等によって、<mark>効率的なコード生成</mark>に必要なプログラミング技法

危険な例

生成される危険なコード

< for
$$x'=0$$
 to n do
$$a[x',y'] \leftarrow \text{complex calculation} \\ b[x',y'] \leftarrow \text{complex calculation} >$$

生成される危険なコード

```
< let w = \text{complex calculation in}

for x' = 0 to n do

let u = \text{complex calculation in}

for y' = 0 to m do

a[x', y'] \leftarrow u

b[x', y'] \leftarrow w
```

生成される危険なコード

```
< let w = \text{complex calculation in} - w が x にも y にも依存する式 for <math>x' = 0 to n do let u = \text{complex calculation in} - u が y に依存する式 for <math>y' = 0 to m do a[x',y'] \leftarrow u b[x',y'] \leftarrow w >
```

生成される危険なコード

```
< let w = \text{complex calculation in} - w が x にも y にも依存する式  for x' = 0 to n do let u = \text{complex calculation in} - u が y に依存する式  for y' = 0 to m do a[x',y'] \leftarrow u b[x',y'] \leftarrow w >
```

complex calculation によって挿入できる場所が異なる

- 多段階 let 挿入が可能となっても、安全に挿入できる場所と そうでない場所がある
- 安全に let 挿入を行うためにどうすればよいかを考える必要がある

コード生成の利点と課題

利点

• 「保守性・再利用性の高さ」と「実行性能の高さ」の両立

コード生成の利点と課題

利点

• 「保守性・再利用性の高さ」と「実行性能の高さ」の両立

課題

- パラメータに応じて、非常に多数のコードが生成される
- 生成したコードのデバッグが容易ではない
- **⇒ コード生成の前に安全性を保証したい**

アウトライン

- 1 概要
- 2 研究の目的
- 3 研究の内容

研究の目的

表現力と安全性を兼ね備えたコード生成言語の構築

- 表現力: 多段階 let 挿入, メモ化等の技法を表現
- 安全性: 生成されるコードの一定の性質を静的に検査

研究の目的

表現力と安全性を兼ね備えたコード生成言語の構築

- 表現力: 多段階 let 挿入, メモ化等の技法を表現
- 安全性: 生成されるコードの一定の性質を静的に検査

本研究: 簡潔で強力なコントロールオペレータに基づ くコード生成体系の構築

- コントロールオペレータ shift0/reset0 を利用し、let 挿入などのコード生成技法を表現
- 型システムを構築して型安全性を保証

アウトライン

- 1 概要
- 2 研究の目的
- 3 研究の内容

表現力を上げ(コードレベル での多段階let挿入)、安全性 も保証するためにどうすれば よいのか

本研究の手法



まず表現力について

let **挿入の実現方法**

コード生成器

... for
$$x = e1$$
 to $e2$ do
... for $y = e3$ to $e4$ do
set $\(x,y\)$... comp calc

生成されるコード

```
<let u' = \text{comp calc' in}
for x' = e1' to e2' do
for y' = e3' to e4' do
a[x', y'] \leftarrow u'>
```

let **挿入の実現方法**

コード生成器

...
$$\underline{\mathbf{for}}\ x = e1\ \underline{\mathbf{to}}\ e2\ \underline{\mathbf{do}}$$
 ... $\underline{\mathbf{for}}\ y = e3\ \underline{\mathbf{to}}\ e4\ \underline{\mathbf{do}}$ $\underline{\mathbf{set}}\ \(x,y\)$... $\underline{\mathbf{comp}}\ \mathbf{calc}$

生成されるコード

```
<let u' = \text{comp calc' in}
for x' = e1' to e2' do
for y' = e3' to e4' do
a[x', y'] \leftarrow u'>
```

shift0/reset0の導入

… のところに後に説明する shift0/reset0 というコントロールオペレータを用いることで、多段階 let 挿入を行う

```
コード生成器: reset0 for x=e1 to e2 do for y=e3 to e4 do set a (x,y) shift0 k \rightarrowlet u=cc in throw k u
```

```
コード生成器: reset0 for x=e1 to e2 do for y=e3 to e4 do set a (x,y) shift0 k \rightarrowlet u=cc in throw k u k \Leftarrow  for x=e1 to e2 do for y=e3 to e4 do set a (x,y) []
```

```
コード生成器: reset0 for x = e1 to e2 do
                      for y = e3 to e4 do
                       set a(x,y) shift 0 k \rightarrow \text{let } u = cc in throw k u
            k \Leftarrow \text{ for } x = e1 \text{ to } e2 \text{ do}
                      for y = e3 to e4 do
                        set a(x,y)
 生成コード: < let u' = cc' in
                for x' = e1' to e2' do
                  for y' = e3' to e4' do
                   a[x',y'] \leftarrow u' >
```

```
コード生成器: reset0 for x=e1 to e2 do reset0 for y=e3 to e4 do set a (x,y) shift0 k_1 \rightarrow \underline{\text{let}}\ u=cc1 in throw k_1 u set b (x,y) shift0 k_2 \rightarrow \underline{\text{let}}\ w=cc2 in throw k_2 w
```

```
コード生成器: reset0 for x = e1 to e2 do
               reset0 for y = e3 to e4 do
                set a(x,y) shift 0 k_1 \rightarrow \text{let } u = cc1 in throw k_1 u
                set b(x,y) shift 0 k_2 \rightarrow \underline{let} w = cc2 \underline{in} throw k_2 w
 生成コード: < let w' = cc2' in
                for x' = e1' to e2' do
                  let u' = cc1' in
                    for y' = e3' to e4' do
                    a[x',y'] \leftarrow u'
                    b[x',y'] \leftarrow w' >
```

次に安全性

コード生成前の段階で,安全 なコードかどうかを判断する

$$\frac{\mathbf{for} \ \mathbf{x} = \mathbf{e1} \ \underline{\mathbf{to}} \ \mathbf{e2} \ \underline{\mathbf{do}}}{\mathbf{\gamma 1}}$$

$$\frac{\mathbf{for} \ \mathbf{y} = \mathbf{e3} \ \underline{\mathbf{to}} \ \mathbf{e4} \ \underline{\mathbf{do}}}{\mathbf{\gamma 2}}$$

$$\frac{\mathbf{set}}{\mathbf{a} \ (\mathbf{x}, \mathbf{y}) \ \text{complex calc}}$$

スコープ	使えるコード変数
$\gamma 0$	なし
$\gamma 1$	X
$\gamma 2$	x,y

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \ge \gamma 1, \ x : \langle \text{int} \rangle \hat{\ } \gamma 1, \ y : \langle \text{int} \rangle \hat{\ } \gamma 2$$

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \ge \gamma 1, \ x : \langle \text{int} \rangle^{\hat{}} \gamma 1, \ y : \langle \text{int} \rangle^{\hat{}} \gamma 2$$

$$\Gamma \vdash x : \langle \text{int} \rangle^{\hat{}} \gamma 2 \text{ OK}$$

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \ge \gamma 1, \ x : \langle \text{int} \rangle^{\hat{}} \gamma 1, \ y : \langle \text{int} \rangle^{\hat{}} \gamma 2$$

$$\Gamma \vdash x : \langle \text{int} \rangle^{\hat{}} \gamma 2 \text{ OK}$$

$$\Gamma \; \vdash \; y : \langle \mathbf{int} \rangle \, \widehat{} \, \gamma 1 \; \; \mathbf{NG}$$

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \ge \gamma 1, \ x : \langle \text{int} \rangle^{\hat{}} \gamma 1, \ y : \langle \text{int} \rangle^{\hat{}} \gamma 2$$

 $\Gamma \vdash x : \langle \text{int} \rangle^{\hat{}} \gamma 2 \text{ OK}$

 $\Gamma \vdash y : \langle \text{int} \rangle^{\hat{}} \gamma 1 \text{ NG}$

 $\Gamma \vdash x + y : \langle int \rangle^{\gamma} 1 \text{ NG}$

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \ge \gamma 1, \ x : \langle \text{int} \rangle^{\hat{}} \gamma 1, \ y : \langle \text{int} \rangle^{\hat{}} \gamma 2$$

 $\Gamma \vdash x : \langle \text{int} \rangle^{\hat{}} \gamma 2 \text{ OK}$

 $\Gamma \vdash y : \langle \mathsf{int} \rangle^{\hat{}} \gamma 1 \mathsf{NG}$

 $\Gamma \vdash x + y : \langle int \rangle^{\gamma} 1 \text{ NG}$

 $\Gamma \vdash x + y : \langle int \rangle^{2} OK$

型システムでコード変数のスコープを表現:

$$\Gamma = \gamma 2 \geq \gamma 1, \ x: \langle \mathrm{int} \rangle \widehat{\ } \gamma 1, \ y: \langle \mathrm{int} \rangle \widehat{\ } \gamma 2$$

 $\Gamma \; \vdash \; x : \langle \mathsf{int} \rangle {}^{\color{red} \boldsymbol{\gamma}} 2 \; \; {\color{red} \mathsf{OK}}$

 $\Gamma \vdash y : \langle \mathsf{int} \rangle^{\hat{}} \gamma 1 \mathsf{NG}$

 $\Gamma \vdash x + y : \langle int \rangle^{\gamma} 1 \text{ NG}$

 $\Gamma \vdash x+y : \langle int \rangle^{\gamma}$ OK

コードレベルのラムダ抽象の型付け規則で固有変数条件を利用:

$$\frac{\Gamma, \ \gamma_2 \geq \gamma_1, \ x: \langle t_1 \rangle \widehat{\ } \gamma_2 \vdash e: \langle t_2 \rangle \widehat{\ } \gamma_2}{\Gamma \vdash \underline{\lambda} x.e: \langle t_1 \rightarrow t_2 \rangle \widehat{\ } \gamma_1} \ (\gamma_2 \text{ is eigen var})$$

環境識別子(EC)を利用したスコープ表現

先行研究:

- 局所的なスコープをもつ破壊的変数をもつコード生成の体系に対する (型安全な) 型システムの構築
 [Sudo,Kiselyov,Kameyama 2014]
- グローバルなスコープをもつ破壊的変数への拡張 [Kiselyov, Kameyama, Sudo 2016]
- コントロールオペレータには非対応

環境識別子(EC)を利用したスコープ表現

先行研究:

- 局所的なスコープをもつ破壊的変数をもつコード生成の体系に対する (型安全な) 型システムの構築
 [Sudo,Kiselyov,Kameyama 2014]
- グローバルなスコープをもつ破壊的変数への拡張 [Kiselyov, Kameyama, Sudo 2016]
- コントロールオペレータには非対応

問題点:

shift0/reset0 などのコントロールオペレータは、スコープの包含 関係を逆転させてしまう。

環境識別子(EC)の問題点

コード生成前・後でスコープの包含関係が逆転

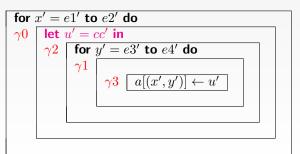
コード生成器:

コード生成前・後でスコープの包含関係が逆転

コード生成器:

```
 \begin{array}{c|c} \underline{\text{for}} \ x = e1 \ \underline{\text{to}} \ e2 \ \underline{\text{do}} \\ \hline \gamma 0 \\ \hline \text{reset0} \ \underline{\text{for}} \ y = e3 \ \underline{\text{to}} \ e4 \ \underline{\text{do}} \\ \hline \gamma 1 \\ \hline \hline \text{shift0} \ k \rightarrow \underline{\text{let}} \ u = cc \ \text{in} \\ \hline \gamma 2 \\ \hline \hline \text{throw} \ k \\ \hline \gamma 3 \\ \hline \hline \text{set} \ a \ (x,y) \ u \\ \hline \end{array}
```

生成コード:



コード生成前・後でスコープの包含関係が逆転

	スコープ	使えるコード変数
	$\gamma 0$	X
コード生成器:	$\gamma 1$	x,y
-	$\gamma 2$	x,y,u
	$\gamma 3$	x,y,u

	スコープ	使えるコード変数
	$\gamma 0$	X
生成コード:	$\gamma 1$	x,u
	$\gamma 2$	x,y,u
	$\gamma 3$	x,y,u

• 生成前と生成後の $\gamma 1$ と $\gamma 2$ の間には順序関係がない

本研究の解決策 ECのジョイン



本研究の解決策 ECのジョイン



- γ_1 のコードレベル変数は γ_2 では使えない
- γ_2 のコードレベル変数は γ_1 では使えない
- γ_1,γ_2 のコードレベル変数は γ_3 で使える

本研究の解決策 ECのジョイン



- γ_1 のコードレベル変数は γ_2 では使えない
- γ_2 のコードレベル変数は γ_1 では使えない
- γ_1,γ_2 のコードレベル変数は γ_3 で使える
- ⇒ Sudo らの体系に ∪ を追加

コード生成+shift0/reset0 の型システム (の一部)

reset0:

$$\frac{\Gamma \vdash e : \langle t \rangle^{\gamma} \; ; \; \langle t \rangle^{\gamma}, \sigma}{\Gamma \vdash \mathbf{reset0} \; e : \langle t \rangle^{\gamma} \; ; \; \sigma}$$

shift0:

$$\frac{\Gamma, \ k: (\langle t_1 \rangle^{\gamma_1} \Rightarrow \langle t_0 \rangle^{\gamma_0}) \sigma \vdash e: \langle t_0 \rangle^{\gamma_0}; \sigma \quad \Gamma \models \gamma_1 \geq \gamma_0}{\Gamma \vdash \mathbf{shift0} \ k.e: \langle t_1 \rangle^{\gamma_1}; \ \langle t_0 \rangle^{\gamma_0}, \sigma}$$

throw:

$$\frac{\Gamma \vdash v : \langle t_1 \rangle^{\gamma_1 \cup \gamma_2}; \sigma \quad \Gamma \models \gamma_2 \ge \gamma_0}{\Gamma, \ k : (\langle t_1 \rangle^{\gamma_1} \Rightarrow \langle t_0 \rangle^{\gamma_0}) \sigma \vdash \mathbf{throw} \ k \ v : \langle t_0 \rangle^{\gamma_2}; \sigma}$$

APPENDIX

アウトライン

4 健全性の証明

健全性の証明 (Subject Reduction)

型安全性 (型システムの健全性; Subject Reduction 等の性質) を厳密に証明する.

Subject Redcution Property

 $\Gamma \vdash M : \tau$ が導ければ (プログラム M が型検査を通れば),M を計算して得られる任意の N に対して, $\Gamma \vdash N : \tau$ が導ける (N も型検査を通り,M と同じ型,同じ自由変数を持つ)