安全なコード移動が可能な
コード生成言語の型システムの設計と実装

大石純平
指導教員 亀山幸義

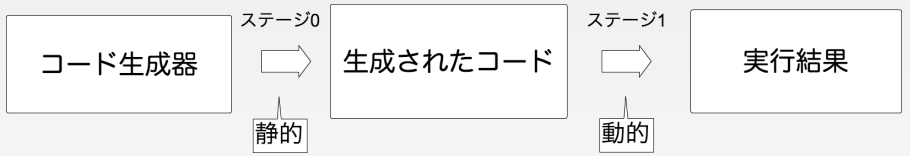
筑波大学 コンピュータサイエンス専攻

2017/1/27
筑波大学修論審査会

アウトライン

アウトライン

目的

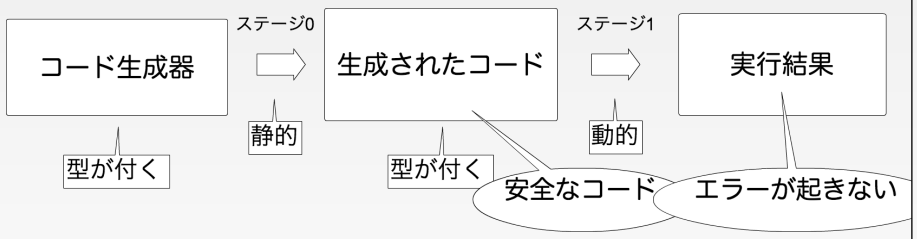


- コード生成をサポートするプログラム言語
  (=コード生成言語)

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

コード生成前に型付け，生成後のコードの型安全性を保証



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，多段階 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

コード生成言語による記述例

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

let 挿入 (コード移動) の実現方法

コード生成器

$\underline{\textbf{for}}\ x = e1\ \underline{\textbf{to}}\ e2\ \underline{\textbf{do}}$
$\underline{\textbf{for}}\ y = e3\ \underline{\textbf{to}}\ e4\ \underline{\textbf{do}}$
$\underline{\textbf{set}}$ <a> (x, y)
$\underline{\textbf{let}}\ u\ =$ cc $\underline{\textbf{in}}$ u

$\rightsquigarrow^*$

生成したいコード

<for $x' = e1'$ to $e2'$ do
let u' = cc$'$ in
for $y' = e3'$ to $e4'$ do
$a[x', y'] \leftarrow u'$>

shift0/reset0 の導入

shift0/reset0 等を用いることで，(多段階)let 挿入等を行う

アウトライン

コード生成前後でコードが移動する

コード生成器		生成されるコード
for $x = e1$ to $e2$ do		< for $x' = e1'$ to $e2'$ do
reset0 for $y = e3$ to $e4$ do	$\rightsquigarrow^*$	let $u' = y'$ in
shift0 $k \rightarrow$ let $u = y$ in		for $y' = e3'$ to $e4'$ do
throw k (set a (x, y) u)		$a[(x', y')] \leftarrow u'$ >

Scope Extrusion

(コード移動により) 意図した束縛から，変数が抜け出てしまうこと

アウトライン

解決策

環境識別子 (EC) を利用したスコープ表現 [Sudo+2014]

$\gamma 0$ **for** $x = e1$ **to** $e2$ **do**
$\gamma 1$ **for** $y = e3$ **to** $e4$ **do**
$\gamma 2$ **set** $a\ (x, y)\ cc$

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

$$\gamma 2 \geq \gamma 1 \geq \gamma 0$$

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.

コード生成+shift0/reset0 の型システム (の一部)

reset0:

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

shift0:

$$\frac{\Gamma,\ k : \langle t1 \rangle\hat{}\gamma 1 \Rightarrow \langle t0 \rangle\hat{}\gamma 0 \vdash e : \langle t0 \rangle\hat{}\gamma 0 \ ; \ \sigma \quad \Gamma \models \gamma 1 \geq \gamma 0}{\Gamma \vdash \textbf{shift0}\ k \rightarrow e : \langle t1 \rangle\hat{}\gamma 1 \ ; \ \langle t0 \rangle\hat{}\gamma 0, \sigma}$$

throw:

$$\frac{\Gamma \vdash v : \langle t1 \rangle\hat{}\gamma 1 \cup \gamma 2 \ ; \ \sigma \quad \Gamma \models \gamma 2 \geq \gamma 0}{\Gamma,\ k : \langle t1 \rangle\hat{}\gamma 1 \Rightarrow \langle t0 \rangle\hat{}\gamma 0 \vdash \textbf{throw}\ k\ v : \langle t0 \rangle\hat{}\gamma 2 \ ; \ \sigma}$$

型付けの例 (1)

$e =$ **reset0** **for** x = e1 **to** e2 **do** **shift0** k → **let** u= **int** 3 x **in** **throw** k

使用できるコード変数 $\gamma 0 = \{\},\ \gamma 1 = \{x\},\ \gamma 2 = \{u\},\ \gamma 1 \cup \gamma 2 = \{u, x\}$

$$\dfrac{\dfrac{\dfrac{\overline{\Gamma b \vdash u : \langle t \rangle\hat{}\gamma 1 \cup \gamma 2;\ \sigma} \quad \vdots}{\Gamma b \vdash \textbf{throw}\ k\ u : \langle t \rangle\hat{}\gamma 2;\ \epsilon \quad \Gamma a \vdash \textbf{int}\ 3\ x : \langle t \rangle\hat{}\gamma 0;\ \epsilon}}{\Gamma a \vdash \textbf{let}\ u = \ldots : \langle t \rangle\hat{}\gamma 0;\ \epsilon}\ (\gamma 2^*)}{\dfrac{\gamma 1 \geq \gamma 0,\ x : \langle t \rangle\hat{}\gamma 1 \vdash \textbf{shift0}\ k\ \rightarrow\ \ldots : \langle t \rangle\hat{}\gamma 1;\ \langle t \rangle\hat{}\gamma 0}{\vdash \textbf{for}\ x = \ldots : \langle t \rangle\hat{}\gamma 0;\ \langle t \rangle\hat{}\gamma 0}\ (\gamma 1^*)}$$
$$\vdash e : \langle t \rangle\hat{}\gamma 0;\ \epsilon$$

$$\Gamma a = \gamma 1 \geq \gamma 0,\ x : \langle t \rangle\hat{}\gamma 1,\ k : \langle t \rangle\hat{}\gamma 1 \Rightarrow \langle t \rangle\hat{}\gamma 0$$
$$\Gamma b = \Gamma a,\ \gamma 2 \geq \gamma 0,\ u : \langle t \rangle\hat{}\gamma 2$$

型付けの例 (2)

$e' =$ **reset0** (**for** $x = e1$ **to** $e2$ **do** **reset0** (**for** $y = e3$ **to** $e4$ **do**

shift0 $k_2 \to$ **shift0** $k_1 \to$ **let** $u =$ $x\ y$ **in** **throw** k_1 **reset0** (**throw** k_2 $e5$

$$\dfrac{\dfrac{\dfrac{\Gamma e \vdash\ e5 : \langle t\rangle\hat{}\gamma 2 \cup \gamma 1 \cup \gamma 3;\ \ \epsilon}{\Gamma e \vdash \textbf{throw}\ k_2\ e5 : \langle t\rangle\hat{}\gamma 1 \cup \gamma 3;\ \langle t\rangle\hat{}\gamma 1 \cup \gamma 3}}{\Gamma e \vdash \textbf{reset0}\ \textbf{throw}\ k_2\ e5 : \langle t\rangle\hat{}\gamma 1 \cup \gamma 3;\ \epsilon} \qquad \vdots}{\dfrac{\Gamma e = \Gamma d, \gamma 3 \geq \gamma 0, u : \langle t\rangle\hat{}\gamma 3 \vdash \textbf{throw}\ k_1\ \ldots : \langle t\rangle\hat{}\gamma 3;\ \epsilon \quad \Gamma d \vdash \boxed{x\ y} : \langle t\rangle\hat{}\gamma 0; \epsilon}{\dfrac{\Gamma d = \Gamma c, k_1 : \langle t\rangle\hat{}\gamma 1 \Rightarrow \langle t\rangle\hat{}\gamma 0 \vdash \textbf{let}\ u = \ldots : \langle t\rangle\hat{}\gamma 0;\ \epsilon}{\dfrac{\Gamma c = \Gamma b, k_2 : \langle t\rangle\hat{}\gamma 2 \Rightarrow \langle t\rangle\hat{}\gamma 1 \vdash \textbf{shift0}\ k_1 \ldots : \langle t\rangle\hat{}\gamma 1;\quad \langle t\rangle\hat{}\gamma 0}{\dfrac{\Gamma b = \Gamma a, \gamma 2 \geq \gamma 1,\ y : \langle t\rangle\hat{}\gamma 2 \vdash \textbf{shift0}\ k_2 \ldots : \langle t\rangle\hat{}\gamma 2;\quad \langle t\rangle\hat{}\gamma 1, \langle t\rangle\hat{}\gamma 0}{\dfrac{\Gamma a \vdash \textbf{for}\ y = \ldots : \langle t\rangle\hat{}\gamma 1;\quad \langle t\rangle\hat{}\gamma 1, \langle t\rangle\hat{}\gamma 0}{\dfrac{\Gamma a = \gamma 1 \geq \gamma 0,\ x : \langle t\rangle\hat{}\gamma 1 \vdash \textbf{reset0} \cdots : \langle t\rangle\hat{}\gamma 1;\quad \langle t\rangle\hat{}\gamma 0}{\dfrac{\vdash \textbf{for}\ x = \ldots : \langle t\rangle\hat{}\gamma 0;\quad \langle t\rangle\hat{}\gamma 0}{\vdash e' = \textbf{reset0} \cdots : \langle t\rangle\hat{}\gamma 0;\quad \epsilon}}}}}}}$$

$(\gamma 2^*)$ $(\gamma 1^*)$

$\Gamma d = \cdots,\ x : \langle t\rangle\hat{}\gamma 1,\ y : \langle t\rangle\hat{}\gamma 2,\ \gamma 1 \geq \gamma 0,\ \gamma 2 \geq \gamma 1,\ \cdots$

型推論アルゴリズム

型推論アルゴリズム

$\Gamma,\ L,\ \sigma,\ t,\ e$ が与えられたとき，$\Theta(\Gamma \vdash^L e : t;\ \sigma)$ が成立するような代入 Θ があるかどうか判定する

制約生成

与えられた項に対して，型，EC，エフェクトに関する制約を返す

制約解消

その得られた制約を解消し，その制約を満たす代入 Θ または fail を返す

制約生成

制約生成用の型システム T_2

T_1

$$\dfrac{\Gamma \vdash u : \langle \mathtt{int}\rangle^{\gamma}; \sigma \quad \Gamma \vdash w : \langle \mathtt{int}\rangle^{\gamma}; \sigma}{\Gamma \vdash u \pm w : \langle \mathtt{int}\rangle^{\gamma}; \sigma} + \dfrac{\Gamma \vdash e : \langle t\rangle^{\gamma_1}; \sigma \quad \Gamma \models \gamma_2 \geq \gamma_1}{\Gamma \vdash e : \langle t\rangle^{\gamma_2}; \sigma}$$

T_2

$$\dfrac{\Gamma \vdash u : \langle \mathtt{int}\rangle^{\gamma}; \sigma \quad \Gamma \vdash w : \langle \mathtt{int}\rangle^{\gamma}; \sigma}{\Gamma \vdash u \pm w : t; \sigma}\ Constr;\ \Gamma \models t \geq \langle \mathtt{int}\rangle^{\gamma}$$

- 下から上に一意的に適用
- 規則適用時に制約を生成

型に関する順序 $t_1 \geq t_2$

コード型か普通の型か判断できないため，型に関する順序 $\geq$ の導入を行った

制約解消

制約 C

型	$t0 = t1$	$t0 \geq t1$
EC	$\gamma 0 = \gamma 1$	$\gamma 0 \geq \gamma 1 \gamma 0 \geq \gamma 1$
エフェクト (型の列)	$\sigma 0 = \sigma 1$	

制約に対する解の存在判定

型， EC，**エフェクトに対する単一化等をおこなう**

ここでは， EC **の不等式制約の解消について説明をする**

制約解消：EC の不等式制約の解消の例

$$\underline{\lambda} x.(\underline{\textbf{int}}\ 1\ \underline{+}\ x) : t1\langle \texttt{int} \rightarrow \texttt{int}\rangle\hat{}\,d0$$

$$C = \{ \qquad\qquad\qquad\qquad \}$$

$$\Theta = \{t1 := \langle t2 \rightarrow t3\rangle\hat{}\,\gamma',\ t2 := \texttt{int},\ t3 := \texttt{int},\ t4 := \texttt{int},$$
$$\gamma' := d0,\ \gamma 1 := d0,\ \gamma 2 := d1,\ \gamma 3 := d0\}$$

EC の変数の除去を行う

$\gamma 1$ **を選ぶ** C **から** $\gamma' \geq \gamma 1$ **を消去し** Θ **に代入** $\gamma 1 := d0$ **を追加**

$\gamma 2$ **を選ぶ** C **から** $\gamma 2 \geq d1$ **を消去し** Θ **に代入** $\gamma 2 := d1$ **を追加**

$\gamma 3$ **を選ぶ** C **から** $\gamma 2 \geq \gamma 3$ **を消去し** Θ **に代入** $\gamma 3 := d0$ **を追加**

研究成果

コード生成 + shift0/reset0 の体系に対する

- **型システムの設計**
- **型推論アルゴリズムの開発**

アウトライン

まとめと今後の課題

まとめ

- コード生成言語にコード移動を許す仕組み (shift0/reset0) を導入し，その安全性を保証するための型システムの設計を行い
  - 安全性：Scope extrusion が起きないようにする
- 型推論アルゴリズムの開発を行った (実装については制約生成まで)

今後の課題

- 設計した型システムの健全性の証明（Subject reduction）の完成
- 型推論アルゴリズム (制約解消) の実装の完成

APPENDIX

アウトライン

shift/reset と shift0/reset0 の意味論

shift/reset

$$\textbf{reset}\ (E[\textbf{shift}\ k \to e]) \rightsquigarrow \textbf{reset}\ e\{k \Leftarrow E\}$$
$$\textbf{reset}\ (E[\textbf{shift}\ k \to e]) \rightsquigarrow \textbf{reset}\ e[k := \lambda x.\textbf{reset}\ E[x]]$$

shift0/reset0

$$\textbf{reset0}\ (E[\textbf{shift0}\ k \to e]) \rightsquigarrow e\{k \Leftarrow E\}$$
$$\textbf{reset0}\ (E[\textbf{shift0}\ k \to e]) \rightsquigarrow e[k := \lambda x.\textbf{reset0}\ E[x]]$$

s0/r0 を使う理由

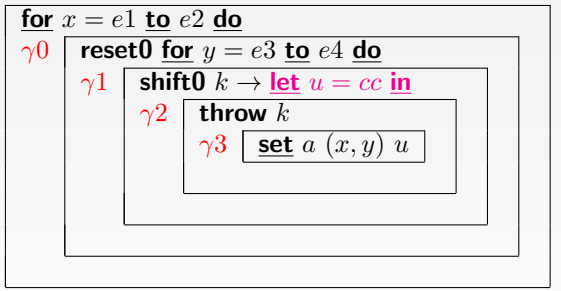
reset0 は評価後に reset0 が付かないので，/shift0/reset0 をネストさせたときに，複数の reset0 をとびこえることができるので，今回の主目的である多段階 let 挿入が行える

アウトライン

コード生成前・後でスコープの包含関係が逆転

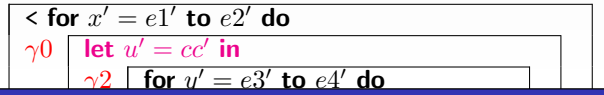
$\gamma 3 \geq \gamma 2 \geq \gamma$

コード生成器:

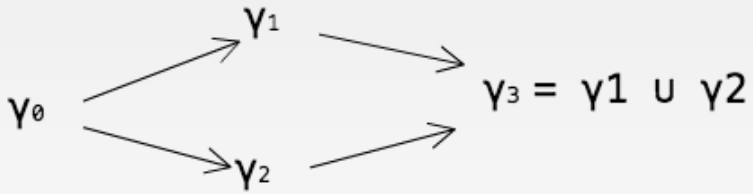


使用できるコー

$\gamma 0 = \{x\}$,
$\gamma 1 = \{x, y\}$,
$\gamma 2 = \{u, x\}$,
$\gamma 3 = \{u, x, y\}$



Union



- $\gamma 1$ のコード変数は $\gamma 2$ では使ってはいけない
- $\gamma 2$ のコード変数は $\gamma 1$ では使ってはいけない

$\Rightarrow$ $\gamma 1$ と $\gamma 2$ の間に順序を付けない

- $\gamma 1, \gamma 2$ のコード変数は $\gamma 3$ で使ってよい

$\Rightarrow$ Sudo らの体系に $\cup$ (ユニオン) を追加

アウトライン

型の単一化の例

$$y \to (\texttt{int} \to w) \to x$$
$$\stackrel{?}{=} (x \to z) \to (x \to z)$$

単一化問題

両辺が同じ型になるような型変数への代入が存在するかどうかを判定することである.

$$\Theta = \{x := \texttt{int} \to w,\ y := (\texttt{int} \to w) \to (\texttt{int} \to w),$$
$$z := \texttt{int} \to w\}$$

$$(\texttt{int} \to w) \to (\texttt{int} \to w) \to (\texttt{int} \to w) \to (\texttt{int} \to w)$$
$$= (\texttt{int} \to w) \to (\texttt{int} \to w) \to (\texttt{int} \to w) \to (\texttt{int} \to w)$$

型に関する順序

$$\frac{\Gamma \vdash e : \langle t \rangle \hat{}\, \gamma 1; \sigma \quad \Gamma \models \gamma 2 \geq \gamma 1}{\Gamma \vdash e : \langle t \rangle \hat{}\, \gamma 2; \sigma} \qquad \frac{\Gamma \vdash e : t; \sigma}{\Gamma \vdash e : t; \sigma}$$

元の型システム T_1 ではこの普通の型，コード型それぞれを一つで表すような表現はなかったため，T_2 ではこのような

$$\frac{\Gamma \vdash e : t2; \sigma}{\Gamma \vdash e : t1; \sigma}\ Constr;\ \Gamma \models t1 \geq t2$$

型に関する順序を導入して一つで表せるようにしている規則を適用するたびに分岐が増えていくために一つの表現で表している

アウトライン

アウトライン

型システムでコード変数のスコープを表現

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

$\gamma 1$	$\gamma 2$
$\Gamma \vdash x : \langle \mathbf{int} \rangle \hat{}\, \gamma 1$ OK	$\Gamma \vdash x : \langle \mathbf{int} \rangle \hat{}\, \gamma 2$ OK
$\Gamma \vdash y : \langle \mathbf{int} \rangle \hat{}\, \gamma 1$ NG	$\Gamma \vdash y : \langle \mathbf{int} \rangle \hat{}\, \gamma 2$ OK
$\Gamma \vdash \underline{x+y} : \langle \mathbf{int} \rangle \hat{}\, \gamma 1$ NG	$\Gamma \vdash \underline{x+y} : \langle \mathbf{int} \rangle \hat{}\, \gamma 2$ 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:

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

この「y は $\gamma 1$ スコープのコード変数としては使えない」という条件を，Sudo らは，自然演繹の固有変数条件を用いて表現した．この型付け規則で，$\gamma 2$ が固有変数であり，規則の結論に $\gamma 2$ が出現してはいけないこと，いいかえれば，下から上に見たとき，$\gamma 2$ が fresh な変数であるという条件となる

アウトライン

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)

EC がなければ，既存手法で証明ができる．今回の場合をカバーするギリギリの型システムを試行錯誤で模索して作成した．あまり強力だと型推論が大変弱いと書きたいものがかけないそこを何度もやりなおしたのでそこの精密な検証ができていない．型推論ができるような型システムのギリギリのところを作成した

answertype の subsumption 規則を reset0 に付けるか throw に付けるか色々と模索を行って型システムを直前まで型推論ができるようにしていたので，そこの精密な検証ができていない