

多段階 let 挿入を行うコード生成言語の 設計

大石純平

筑波大学 大学院
プログラム論理研究室

2016/7/12

アウトライン

- ① 概要
- ② 研究の背景
- ③ 研究の目的
- ④ 研究の内容
- ⑤ まとめと今後

アウトライン

- ① 概要
- ② 研究の背景
- ③ 研究の目的
- ④ 研究の内容
- ⑤ まとめと今後

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

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

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

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

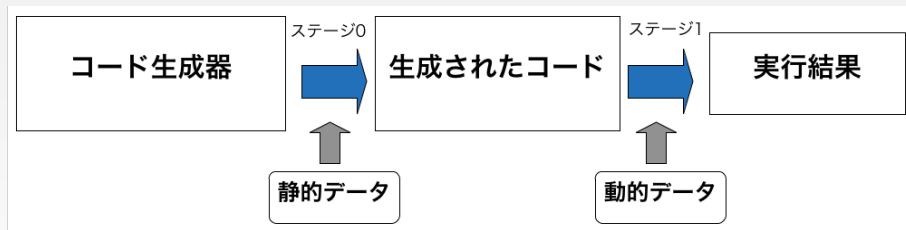
- 効率的なコードの生成
- 安全性の保証

⇒ 多段階 let 挿入を安全に扱うための型システムを構築

アウトライン

- ① 概要
- ② 研究の背景**
- ③ 研究の目的
- ④ 研究の内容
- ⑤ まとめと今後

段階的計算 (Staged Computation)



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

power 関数のコード化

```
power  $x$   $n$  =  $x$                                 if  $n = 1$   
               $x * \text{power } x (n - 1)$            if  $n > 1$ 
```

power 関数のコード化

$\text{power } x \ n$	$= x$	if $n = 1$
	$x * \text{power } x \ (n - 1)$	if $n > 1$

8 に特化したコードの生成を行う

$\text{gen_power } x \ 8 = x * x * x * x * x * x * x * x$

power 関数のコード化

$\text{power } x \ n = x$	if $n = 1$
$x * \text{power } x \ (n - 1)$	if $n > 1$

8 に特化したコードの生成を行う

$\text{gen_power } x \ 8 = x * x * x * x * x * x * x * x$

$\text{gen_power } x \ 8$ は $\text{power } x \ 8$ より高速

- 関数呼び出しがない
- 条件式がない

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

- コード生成プログラムが、安全なコードのみを生成する事を保証
- 安全なコード: 構文, 型, 変数束縛が正しいプログラム

- コード生成プログラムが、安全なコードのみを生成する事を保証
- 安全なコード: 構文, 型, 変数束縛が正しいプログラム

しかし **多段階 let 挿入**等を実現する**計算エフェクト**を含む場合のコード生成の安全性保証は研究途上

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コード移動を許す仕組み
- ループ不変式の移動によって、効率的なコード生成に必要なプログラミング技法

多段階 let 挿入の例

```
for  $i = 0$  to  $n$  in  
  for  $j = 0$  to  $m$  in  
     $y = t$   
     $a[i][j] = b[i] + y$ 
```

多段階 let 挿入の例

```
for  $i = 0$  to  $n$  in  
  for  $j = 0$  to  $m$  in  
     $y = t$   
     $a[i][j] = b[i] + y$ 
```

⇓

```
for  $i = 0$  to  $n$  in  
  for  $j = 0$  to  $m$  in  
    let  $y = t$  in  
     $a[i][j] = b[i] + y$ 
```

— t が i, j に依存した式

多段階 let 挿入の例

```
for  $i = 0$  to  $n$  in  
  for  $j = 0$  to  $m$  in  
     $y = t$   
     $a[i][j] = b[i] + y$ 
```

⇓

```
for  $i = 0$  to  $n$  in
```

```
  let  $y = t$  in
```

```
    for  $j = 0$  to  $m$  in
```

```
       $a[i][j] = b[i] + y$ 
```

— t が i にのみ依存し j には依存しない式

多段階 let 挿入の例

```
for  $i = 0$  to  $n$  in  
  for  $j = 0$  to  $m$  in  
     $y = t$   
     $a[i][j] = b[i] + y$ 
```

⇓

```
let  $y = t$  in      —  $t$  が  $i$  にも  $j$  にも依存しない式  
  for  $i = 0$  to  $n$  in  
    for  $j = 0$  to  $m$  in  
       $a[i][j] = b[i] + y$ 
```

プログラミング言語におけるプログラムを制御するプリミティブ

- exception (例外): C++, Java, ML
- call/cc (第一級継続): Scheme, SML/NJ
- shift/reset (限定継続): Racket, Scala, OCaml
 - 1989 年以降多数研究がある
 - コード生成における let 挿入が実現可能
- shift0/reset0
 - 2011 年以降研究が活発化.
 - コード生成における多段階 let 挿入が可能

アウトライン

- ① 概要
- ② 研究の背景
- ③ 研究の目的**
- ④ 研究の内容
- ⑤ まとめと今後

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 研究の背景
- ③ 研究の目的
- ④ 研究の内容**
- ⑤ まとめと今後

行うこと

- 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行える言語の設計
- その `shift0/reset0` を持つコード生成言語の型システムの構築

困難・問題点, 解決方法

- **shift0/reset0** は shift/reset より強力であるため, 型システムが非常に複雑
- **コード生成言語**の型システムも一定の複雑さを持つ

⇒ 単純な融合は困難

困難・問題点, 解決方法

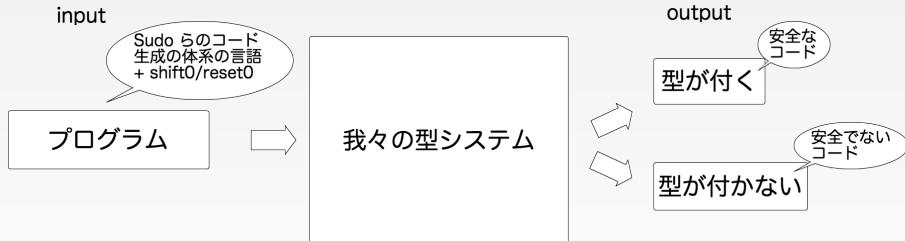
- **shift0/reset0** は shift/reset より強力であるため, 型システムが非常に複雑
- **コード生成言語**の型システムも一定の複雑さを持つ

⇒ 単純な融合は困難

解決方法

コード生成言語の型システムを拡張し, shift0/reset0 の型システムを構築する

本研究の手法



コード生成前

```
cfor  $i = 0$  to  $n$  in  
  cfor  $j = 0$  to  $m$  in  
    clet  $y = t$  in  
       $(a[i][j] = b[i] + y)$ 
```

コード生成前

```
reset0 cfor  $i = 0$  to  $n$  in  
  reset0 cfor  $j = 0$  to  $m$  in  
    shift0  $k_2 \rightarrow$  shift0  $k_1 \rightarrow$  clet  $y = t$  in  
      throw  $k_1$  (throw  $k_2$  ( $a[i][j] = b[i] + y$ ))
```


shift0/reset0 による多段階 let 挿入

コード生成前

```
reset0 cfor  $i = 0$  to  $n$  in  
  reset0 cfor  $j = 0$  to  $m$  in  
    shift0  $k_2 \rightarrow$  shift0  $k_1 \rightarrow$  clet  $y = t$  in  
      throw  $k_1$  (throw  $k_2$  ( $a[\underline{i}][\underline{j}] = b[\underline{i}] + y$ ))
```

```
 $k_1 =$  cfor  $i = 0$  to  $n$  in  
 $k_2 =$  cfor  $j = 0$  to  $m$  in
```

コード生成前

clet $y = t$ in

throw k_1 (throw k_2 ($a[i][j] = b[i] + y$))

$k_1 =$ cfor $i = 0$ to n in

$k_2 =$ cfor $j = 0$ to m in

コード生成前

```
clet  $y = t$  in  
  cfor  $i = 0$  to  $n$  in  
    cfor  $j = 0$  to  $m$  in  
       $(a[i][j] = b[i] + y)$ 
```

型が付く例/付かない例

コード生成前

```
e = reset0 cfor i = 0 to n in  
  reset0 cfor j = 0 to m in  
    shift0 k2 → shift0 k1 → clet y = t in  
      throw k1 (throw k2 (a[i][j] = b[i] + y))
```

型が付く例/付かない例

コード生成前

$e =$ reset0 cfor $i = 0$ to n in
reset0 cfor $j = 0$ to m in
shift0 $k_2 \rightarrow$ shift0 $k_1 \rightarrow$ clet $y = t$ in
throw k_1 (throw k_2 ($a[\underline{i}][\underline{j}] = b[\underline{i}] + y$))

$e \rightsquigarrow^* \text{clet } y = t \text{ in}$
cfor $i = 0$ to n in
cfor $j = 0$ to m in
 $(a[\underline{i}][\underline{j}] = b[\underline{i}] + y)$

型が付く例/付かない例

コード生成前

$e =$ **reset0** cfor $i = 0$ to n in
reset0 cfor $j = 0$ to m in
shift0 $k_2 \rightarrow$ **shift0** $k_1 \rightarrow$ **clet** $y = t$ in
throw k_1 (**throw** k_2 ($a[\underline{i}][\underline{j}] = b[\underline{i}] + y$))

$e \rightsquigarrow^* \text{clet } y = t \text{ in}$
cfor $i = 0$ to n in
cfor $j = 0$ to m in
 $(a[\underline{i}][\underline{j}] = b[\underline{i}] + y)$

$t = \%7$ のとき e は型が付く

$t = a[\underline{i}][\underline{j}]$ や $t = b[\underline{j}]$ のとき e は型が付かない

**安全なコードにのみ型をつける
にはどうすればよいか**

スコープ変数の利用

- γ はスコープ変数を表す.
- ある範囲で使える **自由変数の集合** と思ってもらえば良い.
- γ には, 包含関係があり, それを $\gamma_1 \geq \gamma_0$ というような順序で表している.
- 直感的には γ_0 より γ_1 のほうが使える自由変数が多いという意味である.

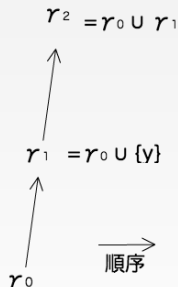
$(\underline{\text{reset0}} \dots^{\gamma_0} (\underline{\text{shift0}} \ k \rightarrow \dots^{\gamma_1} (\underline{\text{throw}} \ k \dots^{\gamma_3})))$

let 挿入が型安全性を保つ条件

従来の場合

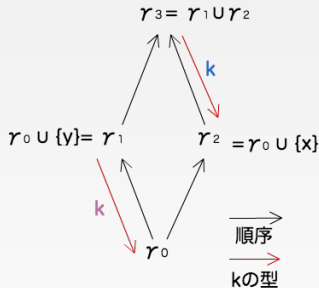
clet $y = \dots^{\gamma_0}$ in \dots^{γ_1}

clet $x = \dots^{\gamma_1}$ in \dots^{γ_2}



shift0/reset0 による let 挿入が型安全性を保つ条件

$(\underline{\text{reset0}} \ \underline{\text{clet}} \ y = \dots^{\gamma_0} \ \underline{\text{in}} \ \dots^{\gamma_1}$
 $(\underline{\text{shift0}} \ k \rightarrow \underline{\text{clet}} \ x = \dots^{\gamma_1} \ \underline{\text{in}} \ \dots^{\gamma_2}$
 $(\underline{\text{throw}} \ k \ \dots^{\gamma_3}))))$



$k : (\langle \cdot \rangle^{\gamma_1} \Rightarrow \langle \cdot \rangle^{\gamma_0})$

$k : (\langle \cdot \rangle^{\gamma_3} \Rightarrow \langle \cdot \rangle^{\gamma_2})$

アウトライン

- ① 概要
- ② 研究の背景
- ③ 研究の目的
- ④ 研究の内容
- ⑤ **まとめと今後**

まとめと今後

- コード生成言語の型システムに `shift0/reset0` を組み込んだ型システムの設計を行った
- その型システムによって型が付く場合と付かない場合の例をみた.
- 今後 `answer type modification` に対応した型システムを設計し, (subject reduction 等の) 健全性の証明を行う