

ソフトウェア構成特論 第11回

大学院理工学研究科 電気電子情報工学専攻 篠埜 功

1 はじめに

今回から、単純型付きラムダ計算にいくつかの構文を追加していき、実際の言語に近づけていく。

2 新たな構文を追加する前の計算体系

新たな構文を追加する前の計算体系として、単純型付きラムダ計算にブール式と数式が入っている、以下の計算体系を用いる。今後、以下の計算体系にさまざまな構文を追加していく。

2.1 式、値、型

式を

$$t ::= x \mid \lambda x : T. t \mid t t \mid \text{true} \mid \text{false} \mid \text{if } t \text{ then } t \text{ else } t \mid \\ 0 \mid \text{succ } t \mid \text{pred } t \mid \text{iszero } t$$

と定義し、値を

$$v ::= \lambda x : T. t \mid \text{true} \mid \text{false} \mid nv \\ nv ::= 0 \mid \text{succ } nv$$

と定義し、型を

$$T ::= T \rightarrow T \mid \text{Bool} \mid \text{Nat}$$

と定義し、型環境を

$$\Gamma ::= \emptyset \mid \Gamma, x : T$$

と定義する。

また、上記の式について、置換は以下のように定義される。

$$\begin{aligned}
[x \mapsto s]x &= s \\
[x \mapsto s]y &= y && \text{if } y \neq x \\
[x \mapsto s](\lambda y : T. t_1) &= \lambda y : T. [x \mapsto s]t_1 && \text{if } y \neq x \text{ and } y \notin FV(s) \\
[x \mapsto s](t_1 t_2) &= ([x \mapsto s]t_1)([x \mapsto s]t_2) \\
[x \mapsto s]\text{true} &= \text{true} \\
[x \mapsto s]\text{false} &= \text{true} \\
[x \mapsto s](\text{if } t_1 \text{ then } t_2 \text{ else } t_3) &= \\
&\quad \text{if } [x \mapsto s]t_1 \text{ then } [x \mapsto s]t_2 \text{ else } [x \mapsto s]t_3 \\
[x \mapsto s]0 &= 0 \\
[x \mapsto s](\text{succ } t) &= \text{succ } ([x \mapsto s]t) \\
[x \mapsto s](\text{pred } t) &= \text{pred } ([x \mapsto s]t) \\
[x \mapsto s](\text{iszero } t) &= \text{iszero } ([x \mapsto s]t)
\end{aligned}$$

ただし、第6回の講義資料で説明した通り、ラムダ抽象に対する置換適用時には、必要に応じて置換対象のラムダ抽象式の束縛変数の付け替えを行う。

2.2 評価規則

評価規則は、以下に示す13個の評価規則とする。これは、第5回の資料で提示した算術式に関する10個の評価規則に、第9回の資料で提示したラムダ式に関する評価規則 (E-APP1), (E-APP2), (E-APPABS) を追加したものである。

$$\begin{aligned}
&\frac{}{\text{if true then } t_2 \text{ else } t_3 \rightarrow t_2} \text{ (E-IFTRUE)} \\
&\frac{}{\text{if false then } t_2 \text{ else } t_3 \rightarrow t_3} \text{ (E-IFFALSE)} \\
&\frac{t_1 \rightarrow t'_1}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t'_1 \text{ then } t_2 \text{ else } t_3} \text{ (E-IF)} \\
&\frac{t_1 \rightarrow t'_1}{\text{succ } t_1 \rightarrow \text{succ } t'_1} \text{ (E-SUCC)} \quad \frac{}{\text{pred } 0 \rightarrow 0} \text{ (E-PREDZERO)} \\
&\frac{}{\text{pred } (\text{succ } nv_1) \rightarrow nv_1} \text{ (E-PREDSUCC)} \quad \frac{t_1 \rightarrow t'_1}{\text{pred } t_1 \rightarrow \text{pred } t'_1} \text{ (E-PRED)} \\
&\frac{}{\text{iszero } 0 \rightarrow \text{true}} \text{ (E-ISZEROZERO)} \\
&\frac{}{\text{iszero } (\text{succ } nv_1) \rightarrow \text{false}} \text{ (E-ISZEROSUCC)} \quad \frac{t_1 \rightarrow t'_1}{\text{iszero } t_1 \rightarrow \text{iszero } t'_1} \text{ (E-ISZERO)} \\
&\frac{t_1 \rightarrow t'_1}{t_1 t_2 \rightarrow t'_1 t_2} \text{ (E-APP1)} \quad \frac{t_2 \rightarrow t'_2}{v_1 t_2 \rightarrow v_1 t'_2} \text{ (E-APP2)} \\
&\frac{}{(\lambda x : T_{11}.t_{12}) v_2 \rightarrow [x \mapsto v_2]t_{12}} \text{ (E-APPABS)}
\end{aligned}$$

2.3 型システム

型付け規則は

$$\begin{array}{c}
\frac{x : T \in \Gamma}{\Gamma \vdash x : T} \text{ (T-VAR)} \quad \frac{\Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2} \text{ (T-ABS)} \\
\\
\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1 t_2 : T_{12}} \text{ (T-APP)} \\
\\
\frac{}{\Gamma \vdash \text{true} : \text{Bool}} \text{ (T-TRUE)} \quad \frac{}{\Gamma \vdash \text{false} : \text{Bool}} \text{ (T-FALSE)} \\
\\
\frac{\Gamma \vdash t_1 : \text{Bool} \quad \Gamma \vdash t_2 : T \quad \Gamma \vdash t_3 : T}{\Gamma \vdash \text{if } t_1 \text{ then } t_2 \text{ else } t_3 : T} \text{ (T-IF)} \\
\\
\frac{}{\Gamma \vdash 0 : \text{Nat}} \text{ (T-ZERO)} \quad \frac{\Gamma \vdash t_1 : \text{Nat}}{\Gamma \vdash \text{succ } t_1 : \text{Nat}} \text{ (T-SUCC)} \quad \frac{\Gamma \vdash t_1 : \text{Nat}}{\Gamma \vdash \text{pred } t_1 : \text{Nat}} \text{ (T-PRED)} \\
\\
\frac{\Gamma \vdash t_1 : \text{Nat}}{\Gamma \vdash \text{iszero } t_1 : \text{Bool}} \text{ (T-ISZERO)}
\end{array}$$

と定義する。

3 基本型

プログラミング言語は通常数、真偽値、文字など、構造を持たない値の集合を基本型として提供する。また、基本型の値を操作する演算も提供する。これまでに自然数の型 `Nat` と真偽値の型 `Bool` を用いてきた。その他の型としては、"hello"のような文字列の型 `String` や 3.14159のような浮動小数点数の型 `Float` などを今後用いる。

基本型を表すメタ変数として A を用いる。また、基本型の名前として A, B, C などを用いる。例えば、term $\lambda x:A. x$ は空の型環境で $A \rightarrow A$ 型を持ち、term $\lambda x:B. x$ は空の型環境で $B \rightarrow B$ 型を持ち、term $\lambda f:A \rightarrow A. \lambda x:A. f(f(x))$ は空の型環境で $(A \rightarrow A) \rightarrow (A \rightarrow A)$ 型を持つ。

4 Unit 型

`unit` という、`Unit` 型の値を導入する。つまり、term 集合に `unit` を追加し、値に `unit` を追加し、型に `Unit` を追加する。また、`Unit` 型の term の評価結果は必ず `unit` になる。また、`unit` に関する型付け規則を以下のように定義する。

$$\frac{}{\Gamma \vdash \text{unit} : \text{Unit}} \text{ (T-UNIT)}$$

`unit` は、(通常)副作用のあるプログラムで用いる。前回までに提示した計算体系では副作用のある term はないが、reference cell への代入や画面への出力などが副作用である。副作用が目的で、値が不要な場合でも何らかの値が必要であり、そのような場合のために `unit` が用いられる。C 言語や Java 言語でいえば `void` 型を使う場合がこれに相当する。

置換の定義に `unit` の場合を追加する。

$$[x \mapsto s]\text{unit} = \text{unit}$$

5 逐次式

副作用のある言語では、2つ以上の式を順番に評価するということがよく行われる。そのために、逐次式 $t_1; t_2$ を用いる。この式を評価するときは、まず t_1 を評価し、その評価結果を捨ててから t_2 の評価をし、 t_2 の評価結果を $t_1; t_2$ の評価結果とする。これを規則で書くと以下ようになる。

$$\frac{t_1 \rightarrow t'_1}{t_1; t_2 \rightarrow t'_1; t_2} \text{ (E-SEQ)} \quad \frac{}{\text{unit}; t_2 \rightarrow t_2} \text{ (E-SEQNEXT)}$$

また以下の型付け規則を追加する。

$$\frac{\Gamma \vdash t_1 : \text{Unit} \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash t_1; t_2 : T_2} \text{ (T-SEQ)}$$

また置換の定義に逐次式の場合を追加する。

$$[x \mapsto s](t_1; t_2) = ([x \mapsto s]t_1); ([x \mapsto s]t_2)$$

$t_1; t_2$ と同等のことを、 $(\lambda x : \text{Unit}. t_2)t_1$ (ただし x は t_2 の自由変数ではない変数) でできるので、 $t_1; t_2$ を $(\lambda x : \text{Unit}. t_2)t_1$ の略記として定義してもよい。

定理 1. 単純型付きラムダ計算を Unit 型、逐次式、E-SEQ、E-SEQNEXT、T-SEQ 規則で拡張した体系を λ^E と書く。また、単純型付きラムダ計算を Unit 型で拡張しただけの体系を λ^I と書く。

λ^E の term が与えられたとき、その中の $t_1; t_2$ の形の term を全部 $(\lambda x : \text{Unit}. t_2)t_1$ (ただし x は t_2 の自由変数ではない変数) で置き換えて得られる λ^I の term を返す関数 $e \in \lambda^E \rightarrow \lambda^I$ を *elaboration function* と呼ぶ。このとき、 λ^E の各 term t に対し、

- $t \rightarrow_E t' \implies e(t) \rightarrow_I e(t')$
- $\Gamma \vdash^E t : T \implies \Gamma \vdash^I e(t) : T$

が成り立つ。ここで、 \rightarrow 、 \vdash の添字により、どちらの計算体系の評価関係あるいは型付け関係であるかを示している。

証明. これらの2つの命題について、 t の構造帰納法で示すことができる。 □

この定理により、derived form を使っていいということが分かる。derived form を使う利点は、内部の言語を変えずに表面の構文だけ追加することができるということにあり、これにより、内部言語で成り立つ性質 (Progress や Preservation のような性質) がそのまま成り立つことになる。

Derived form は syntactic sugar (構文糖衣) あるいは syntax sugar とも呼ばれる。Derived form をその定義で展開することを desugaring あるいは elaboration という。

練習問題 1. 以下の式は空の型環境で型を持つが、その型判定およびその導出木を書け。

$$(\text{unit}; \text{unit}); \text{unit}$$

練習問題 2. 以下の式を正規形になるまで評価せよ。その際、各1ステップ評価の導出木も書け。

$$(\text{unit}; \text{unit}); \text{unit}$$

6 Wildcard

もう一つ有用な derived form として、ラムダ抽象の束縛変数部分で wildcard を使うというものがある。ラムダ抽象の束縛変数が本体内で使われない場合に wildcard を使えば、束縛変数が使われないということが一目で分かるようになり、またプログラムを書く際に変数名を考える手間が省ける。束縛変数部分に wildcard を使った式 $\lambda_ : S.t$ は、 t の自由変数ではない任意の変数を x としたとき、 $\lambda x : S.t$ の derived form として定義すればよい。以下では、derived form ではなく、核言語に wildcard を追加する。

まず、式および値に wildcard を追加する。

$$\begin{aligned} t & ::= \dots \mid \lambda_ : T.t \\ v & ::= \dots \mid \lambda_ : T.t \end{aligned}$$

ワイルドカードに関する型付け規則は

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash \lambda_ : S.t : S \rightarrow T} \text{ (T-WILD)}$$

とすればよい。また評価規則は

$$\frac{}{(\lambda_ : S.t) v \rightarrow t} \text{ (E-WILD)}$$

とすればよい。また置換の定義にワイルドカードの場合を追加する。

$$[x \mapsto s](\lambda_ : S.t) = \lambda_ : S.[x \mapsto s]t$$

練習問題 3. term

$$(\lambda_ : \text{Nat.true}) ((\lambda x : \text{Nat.x}) 0)$$

は空の型環境で型を持つが、その型判定の導出木を書け。

練習問題 4. term

$$(\lambda_ : \text{Nat.true}) ((\lambda x : \text{Nat.x}) 0)$$

を評価せよ。各 1 ステップ評価の導出木も書け。

7 Ascription

term に型を書く (ascribe) ことができると便利が良い場合がある。ある term がある型を持つというつもりでプログラムを書いているとき、その型をそこに書くことによって、もしその term がその型を持っていなかったらコンパイル時に型エラーになるので、プログラマーの意図通りになっているかどうかの確認に使える。また、大きな term の型はプログラムを読む人（あるいはプログラムを書いている人）が一目で分からない場合があり、そのとき、型が書いてあればプログラムを読むのに役に立つ。

ここでは、そのための構文として $t \text{ as } T$ と書く。これに関する型付け規則は、

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash t \text{ as } T : T} \text{ (T-ASCRIIBE)}$$

を用いる。また、評価規則は

$$\frac{}{v \text{ as } T \rightarrow v} \text{ (E-ASCRIBE)}$$

$$\frac{t \rightarrow t'}{t \text{ as } T \rightarrow t' \text{ as } T} \text{ (E-ASCRIBE1)}$$

を用いる。また、置換の定義に ascription の場合を追加する。

$$[x \mapsto s](t \text{ as } T) = ([x \mapsto s]t) \text{ as } T$$

練習問題 5. 型判定

$$\emptyset \vdash ((\lambda x : \text{Nat}.x) 0) \text{ as } \text{Nat} : \text{Nat}$$

は成り立つが、これの導出木を書け。

練習問題 6. term

$$((\lambda x : \text{Nat}.x) 0) \text{ as } \text{Nat}$$

を評価せよ。各 1 ステップ評価の導出木も書け。

補足 $t \text{ as } T$ は t の (定理 1 の意味での) derived form ではない。Ascription を追加した後の体系の term に対して、 $t \text{ as } T$ の形の部分式を全部 t に置き換える関数を e としたとき、 $\Gamma \vdash e(t) : T$ であっても $\Gamma \vdash t : T$ が成り立つとは限らない。(プログラマが as の後の型を間違えて書いてしまった場合に成り立たない。) よって、コンパイラで型検査前に $t \text{ as } T$ を t に置き換えることはできない。 $t \rightarrow t'$ なら $e(t) \rightarrow e(t')$ は成り立つ (ことが証明できる) ので、型検査を通過したら $t \text{ as } T$ を t に置き換えてよいが。

8 let 式

大きな式を書くときに、その部分式に名前を付けることができると便利がよい。ほとんどの言語はこのような機構を備えている。ML 系言語では、 $\text{let } x = t_1 \text{ in } t_2$ を評価するときは、まず式 t_1 を評価して、それに x という名前を付けてから、 t_2 を評価する。

let 式の評価規則は、

$$\frac{}{\text{let } x = v_1 \text{ in } t_2 \rightarrow [x \mapsto v_1]t_2} \text{ (E-LETV)}$$

$$\frac{t_1 \rightarrow t'_1}{\text{let } x = t_1 \text{ in } t_2 \rightarrow \text{let } x = t'_1 \text{ in } t_2} \text{ (E-LET)}$$

とする。また、let 式に関する型付け規則は

$$\frac{\Gamma \vdash t_1 : T_1 \quad \Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \text{let } x = t_1 \text{ in } t_2 : T_2} \text{ (T-LET)}$$

とする。また、置換の定義に let 式の場合を追加する。

$$[x \mapsto s](\text{let } x = t_1 \text{ in } t_2) = \text{let } x = ([x \mapsto s]t_1) \text{ in } t_2$$

$$[x \mapsto s](\text{let } y = t_1 \text{ in } t_2) = \text{let } y = ([x \mapsto s]t_1) \text{ in } ([x \mapsto s]t_2) \quad \text{if } x \neq y$$

上記では let 式に対する置換の定義を 2 つに場合分けして書いたが、束縛変数の名前の付け替えを必要な場合に行うようにすると、ラムダ抽象に対する置換と同様、1 つにまとめてよい。

$$[x \mapsto s](\text{let } y = t_1 \text{ in } t_2) = \text{let } y = ([x \mapsto s]t_1) \text{ in } ([x \mapsto s]t_2) \quad \text{if } x \neq y$$

練習問題 7. let 式

let x = true in if x then 0 else succ 0

は空の型環境で型を持つが、その型判定の導出木を書け。

練習問題 8. let 式

let x = true in if x then 0 else succ 0

を評価せよ。各 1 ステップ評価の導出木も書け。

let 式は derived form として定義することもできる。let 式 $\text{let } x = t_1 \text{ in } t_2$ は $(\lambda x : T_1. t_2) t_1$ の略記として定義すればよい。ただ、ラムダ抽象に T_1 という型注釈があり、これは let 式にはない。つまり、desugaring 時に型 T_1 を何らかの手段で得なければならない。これは型検査器 (type checker) から情報を得ればよく、 t_1 の型を T_1 にすればよい。ただし、let 式のこの derived form による定義は逐次式などの場合とは違い、単なる term の変換ではなく、型判定の導出木の変換である。具体的には、

$$\frac{\frac{\vdots}{\Gamma \vdash t_1 : T_1} \quad \frac{\vdots}{\Gamma, x : T_1 \vdash t_2 : T_2}}{\Gamma \vdash \text{let } x = t_1 \text{ in } t_2 : T_2} \text{ (T-LET)}$$

という形の導出木から

$$\frac{\frac{\frac{\vdots}{\Gamma, x : T_1 \vdash t_2 : T_2}}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2} \text{ (T-ABS)} \quad \frac{\vdots}{\Gamma \vdash t_1 : T_1} \text{ (T-APP)}}{\Gamma \vdash (\lambda x : T_1. t_2) t_1 : T_2}$$

という形の導出木への変換である。

参考 let 多相のある言語では let 式は上記のようにラムダ抽象と関数適用により derived form として定義することはできない。

9 対 (pair)

多くの言語では複合データ型を構築する機構を備える。最も簡単なものでは対 (pair)、より一般には tuple がある。対を追加するために、まず、構文定義に対し

$$t ::= \dots \mid \{t, t\} \mid t.1 \mid t.2$$

のように 3 つの構文を追加する。対を中括弧を使って書くのは、集合の表記と紛らわしく、ML 系言語では普通の括弧を使って書く。ML 系言語ではレコード (後述) は中括弧を使って書くので、参考書ではレコードとの関連を強調するために対を中括弧を使って書いている。

また、値については、

$$v ::= \dots \mid \{v, v\}$$

のように、値の対を追加する。また、型の定義に product type を追加する。

$$T ::= \dots \mid T_1 \times T_2$$

また、評価規則は

$$\begin{array}{c} \frac{}{\{v_1, v_2\}.1 \rightarrow v_1} \text{ (E-PAIRBETA1)} \quad \frac{}{\{v_1, v_2\}.2 \rightarrow v_2} \text{ (E-PAIRBETA2)} \\ \\ \frac{t_1 \rightarrow t'_1}{t_1.1 \rightarrow t'_1.1} \text{ (E-PROJ1)} \quad \frac{t_1 \rightarrow t'_1}{t_1.2 \rightarrow t'_1.2} \text{ (E-PROJ2)} \\ \\ \frac{t_1 \rightarrow t'_1}{\{t_1, t_2\} \rightarrow \{t'_1, t_2\}} \text{ (E-PAIR1)} \quad \frac{t_2 \rightarrow t'_2}{\{v_1, t_2\} \rightarrow \{v_1, t'_2\}} \text{ (E-PAIR2)} \end{array}$$

を用いる。また、型付け規則は、

$$\frac{\Gamma \vdash t_1 : T_1 \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash \{t_1, t_2\} : T_1 \times T_2} \text{ (T-PAIR)}$$

$$\frac{\Gamma \vdash t : T_1 \times T_2}{\Gamma \vdash t.1 : T_1} \text{ (T-PROJ1)} \quad \frac{\Gamma \vdash t : T_1 \times T_2}{\Gamma \vdash t.2 : T_2} \text{ (T-PROJ2)}$$

を用いる。また、置換の定義に対と対の projection の場合を追加する。

$$\begin{aligned} [x \mapsto s]\{t_1, t_2\} &= \{[x \mapsto s]t_1, [x \mapsto s]t_2\} \\ [x \mapsto s](t.1) &= ([x \mapsto s]t).1 \\ [x \mapsto s](t.2) &= ([x \mapsto s]t).2 \end{aligned}$$

(E-PAIR1)、(E-PAIR2) 規則により、対の要素は左が先に評価されることになる。例えば、term

$$\{\text{pred } 4, \text{ if true then false else false}\}.1$$

は以下のように評価される。

$$\begin{aligned} &\{\text{pred } 4, \text{ if true then false else false}\}.1 \\ \rightarrow &\{3, \text{ if true then false else false}\}.1 \\ \rightarrow &\{3, \text{ false}\}.1 \\ \rightarrow &3 \end{aligned}$$

ただし、4 は $\text{succ}(\text{succ}(\text{succ}(\text{succ } 0)))$ の略記であり、3 は $\text{succ}(\text{succ}(\text{succ } 0))$ の略記である。上記の各 1 ステップ評価の導出木は黒板で説明する。

練習問題 9. term

$$(\lambda x : \text{Nat} \times \text{Nat}. x.2) \{\text{pred } 4, \text{ pred } 5\}$$

は空の型環境で型を持つ。その型判定とその導出木を書け。

練習問題 10. term

$$(\lambda x : \text{Nat} \times \text{Nat}. x.2) \{\text{pred } 4, \text{ pred } 5\}$$

を評価せよ。各 1 ステップ評価の導出木も書け。