

{title= レコード計算の
多相型システムと型推論 ,

author=



@bd_gfngfn

,

date=

2017年9月30日

}

今回扱う論文

- A. Ohori. ***A Polymorphic record calculus and its compilation***, 1995.
- E. Osinski. ***A Polymorphic Type System and Compilation Scheme for Record Concatenation***, 2006.

今回扱う論文

- A. Ohori. ***A Polymorphic record calculus and its compilation***, 1995.
- E. Osinski. *A Polymorphic Type System and Compilation Scheme for Record Concatenation*, 2006.
(正直なところ間に合いませんでした)

前者に焦点を当てて説明, 後者はごく初歩を紹介

レコードとは

直観的には：

- 各要素にラベルがついたタプル
- いわゆる keyval や dictionary, 連想配列に似ている
(ただしラベル名は文字列値ではなく識別子的に扱う)

$\{\text{foo} = 42, \text{bar} = \text{true}\}$

レコードに対する演算

とりあえず型を度外視するとして、書けると嬉しそうな操作は：

- “第1級な” 構築 (これは書いて当然)

$\text{let } r = \{\text{foo} = 42, \text{bar} = \text{true}\} \text{ in } \dots$

- フィールド取出し (これも当然)

$r.\text{foo} \longrightarrow 42$

- 更新

$r \text{ update } \text{bar } \text{false} \longrightarrow \{\text{foo} = 42, \text{bar} = \text{false}\}$

※ “値を覆い隠している” だけで、破壊的代入ではない

レコードに対する演算

- 拡張 $\{\text{foo} = 42\} \text{ extend bar true}$
 $\longrightarrow \{\text{foo} = 42, \text{bar} = \text{true}\}$

既に含まれるラベルと被る場合, 上書きするか何もしないかで
ヴァリエーションあり

- 接続 (concatenation)

$$\{\text{foo} = 42, \text{bar} = \text{true}\} \mid \& \mid \{\text{foo} = 0, \text{baz} = \text{“オッ”}\}$$
$$\longrightarrow \{\text{foo} = 0, \text{bar} = \text{true}, \text{baz} = \text{“オッ”}\}$$

ラベルが重複する場合を許して後者の値を優先するものと
許さないもののヴァリエーションあり

タプルに比べたレコードのありがたみ

- ラベルのおかげで単純にコードのドキュメント性が向上

タプル流

```
("realDonaldTrump", "Donald J. Trump", true, true)
```

タプルに比べたレコードのありがたみ

- ラベルのおかげで単純にコードのドキュメント性が向上

タプル流

```
("realDonaldTrump", "Donald J. Trump", true, true)
```

レコード流

```
{screen_name= "realDonaldTrump",  
  name= "Donald J. Trump",  
  is_official= true, is_public= true}
```


タプルに比べたレコードのありがたみ

- “多相性”を持ち，そのうえ“組にして扱う要素”の仕様変更に対してもコードが頑強

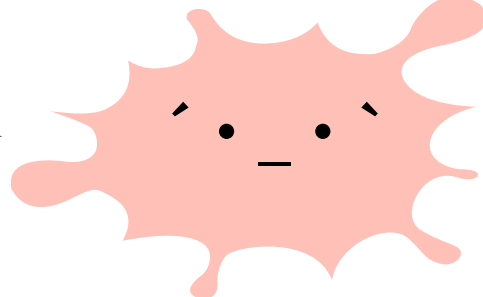
組の要素の
仕様変更のたびに
書き換える必要あり

```
let send_missile = λaccount.  
  let (screen_name, __, __, __) = account in  
  send_reply screen_name “🚀”
```

```
let send_missile = λaccount.  
  send_reply (account.screen_name) “🚀”
```

問題意識

- 型なしだと（当然ながら）取出しなどの際に失敗しうる

`{foo= 42} ■ bar` →  破滅



適切な型システムを設計して，失敗し得ないプログラムに
だけ静的に型がつくようにしたい

ただし，その際に.....

問題意識

- できるだけ保守的になりすぎないようにしたい
(前述の演算のなるべく多くを“自然な範囲で”サポートしたい)
- **主要型** (principal type) が存在し、
かつそれが**型推論**で得られる体系にしたい
- できればより計算上効率的な形式へと
コンパイルできるようにしたい

[Ohori 1995] 概要

- **SML#** の型システムの基礎づけ
- 双対である多相ヴァリアントもサポート（今回は扱わない）
- 一般的な大きい入力に対しても十分高速に型推論が動作
- ✱ レコードの拡張や接続はサポートしない
- ✓ 型推論可能
- ✓ コンパイルによる効率化が可能
- ✓ （主観ながら）推論される型が比較的わかりやすい

[Osinski 2006] 概要

- 博士論文（同著者唯一の publication らしい？）
- ✓ 前述のすべての演算をサポート
- ✓ 型推論可能
- ✓ コンパイルによる効率化が可能
- ▲ （主観ながら）推論される型が複雑でわかりにくい？
- ▲ 実装がまだ無いようで、**実時間の動作に堪えるか不明**
（自分の手で実装して確かめるつもりでしたが間に合わず）

構成

- 背景
- おさらい：Let多相の型推論
- [Ohori 1995]
- [Osinski 2006]
- まとめ

項と型

型註釈なしの項を主として使用

● 値 $v ::= c \mid \lambda x. e$

● 項 $e ::= v \mid x \mid e e \mid \text{let } x = e \text{ in } e$

実用上は不動点 ($\text{fix } x. \lambda x. e$) や条件分岐 $\text{if } e \text{ then } e \text{ else } e$ も必要だが、簡単のため今回は含めない

● 単相型 $\tau ::= b \mid \alpha \mid \tau \rightarrow \tau$

● 多相型 $\sigma ::= \tau \mid \forall \alpha. \sigma$

型付け規則

$$\boxed{\Gamma \vdash e : \tau}$$

型変数の束縛出現の
“インスタンス化”

$$\frac{(\sigma \equiv \Gamma(x), \quad \sigma \geq \tau)}{\Gamma \vdash x : \tau}$$

$$\frac{\Gamma[x \mapsto \tau'] \vdash e : \tau}{\Gamma \vdash (\lambda x. e) : \tau' \rightarrow \tau}$$

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

型変数の量化

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad (\sigma \equiv \text{Gen}(\Gamma, \tau)) \quad \Gamma[x \mapsto \sigma] \vdash e_2 : \tau_2}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau_2}$$

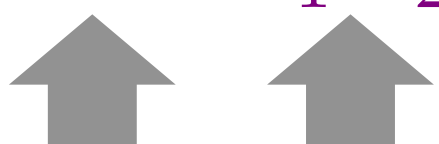
型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$


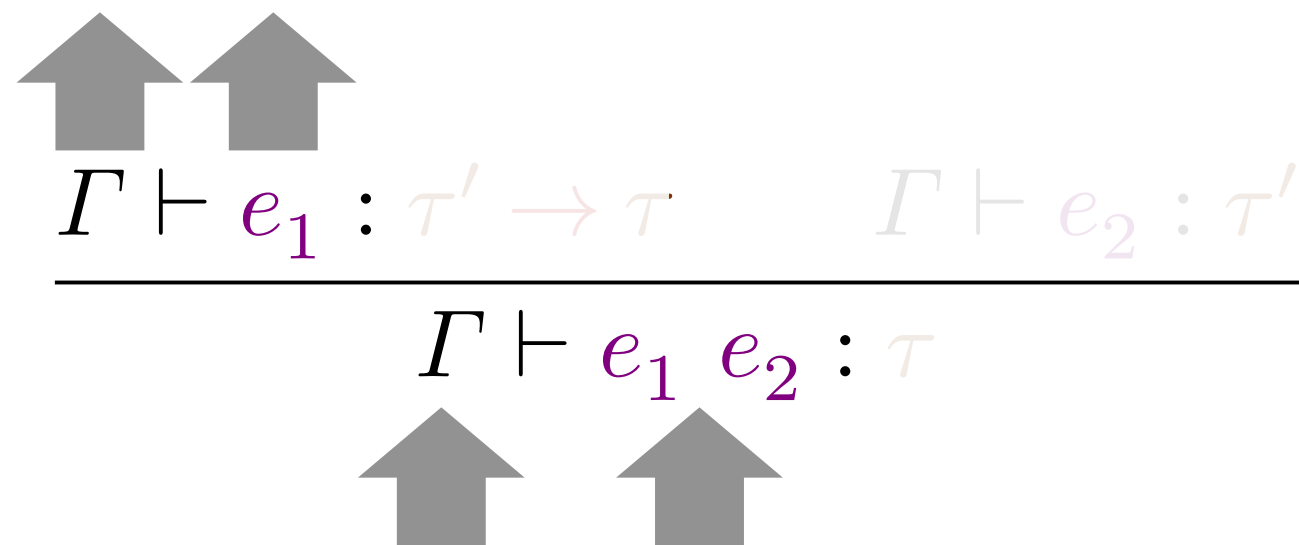
型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”


$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

The diagram illustrates the bottom-up reading of the typing rule. It shows the conclusion $\Gamma \vdash e_1 e_2 : \tau$ at the bottom, with two arrows pointing up to the premises $\Gamma \vdash e_1 : \tau' \rightarrow \tau$ and $\Gamma \vdash e_2 : \tau'$. The premises are written in a lighter color, and the conclusion is in a darker color.

型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

The diagram illustrates the bottom-up reading of the typing rule. It shows the conclusion $\Gamma \vdash e_1 e_2 : \tau$ at the bottom, with two arrows pointing up to the premises $\Gamma \vdash e_1 : \tau' \rightarrow \tau$ and $\Gamma \vdash e_2 : \tau'$. Above each premise, there are two more arrows pointing up, indicating the flow of information from the conclusion back to the assumptions.

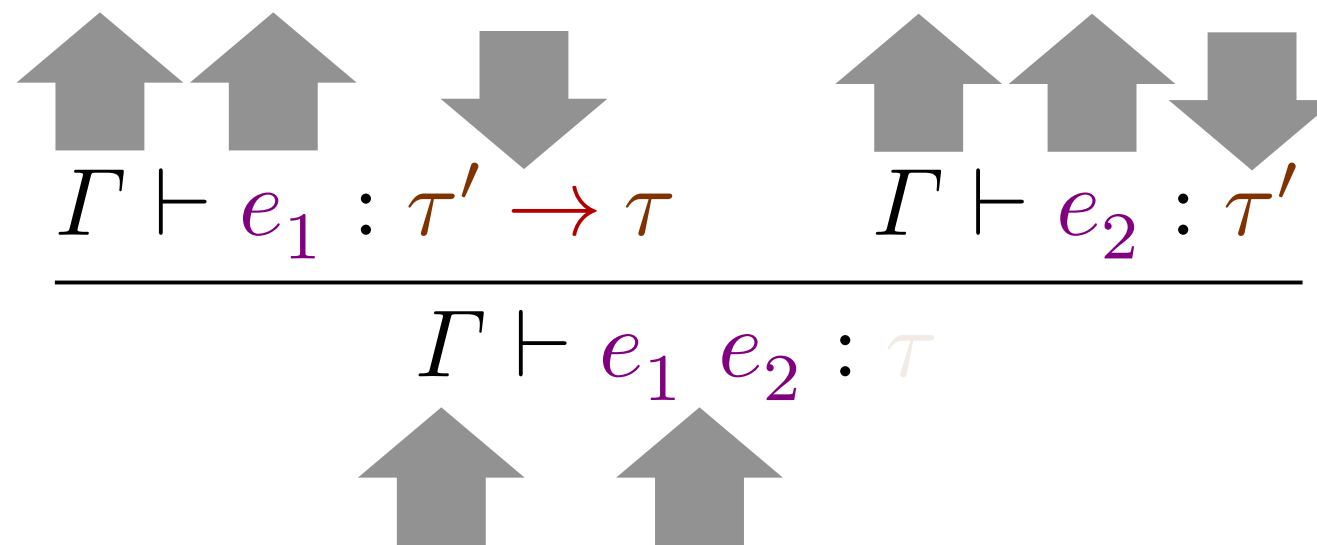
型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”


$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

型付け規則の読み方

1. 純粹に3項関係の帰納的定義として

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

“仮定側”
“帰結側”

2. アルゴリズムとして：“下から読む”

$$\frac{\Gamma \vdash e_1 : \tau' \rightarrow \tau \quad \Gamma \vdash e_2 : \tau'}{\Gamma \vdash e_1 e_2 : \tau}$$

(※不正確, あくまで直観)

型つけ規則から型推論アルゴリズムへ

型推論アルゴリズムとしては、より正確には**型変数**と**代入**（型変数を実際の型へ関連づける部分写像）を使う

つまるところ型推論とは「未知の型を型変数で置いておき、項を走査して種々の制約をもとに代入を育てていく」処理

型が未知のところを
型変数で置いて上る

制約の解消結果として
下りてきた代入

$$\frac{(\alpha \text{ fresh}) \quad \Gamma[x \mapsto \alpha] \vdash e : \tau \downarrow \theta}{\Gamma \vdash (\lambda x. e) : (\theta \alpha \rightarrow \tau) \downarrow \theta}$$

下りてきた代入を
使って実際の型を得る

型つけ規則から型推論アルゴリズムへ

単一化：

型の等式制約を解いて
代入にする処理,
型推論の根幹

$$\frac{\begin{array}{l} \Gamma \vdash e_1 : \tau_1 \downarrow \theta_1 \quad \theta_1 \Gamma \vdash e_2 : \tau_2 \downarrow \theta_2 \\ \alpha \text{ fresh} \quad \theta' := \text{Unify}\left(\left\{\theta_2 \tau_1 \stackrel{?}{=} \tau_2 \rightarrow \alpha\right\}\right), \end{array}}{\Gamma \vdash e_1 e_2 : \theta' \alpha \downarrow \theta' \circ \theta_2 \circ \theta_1}$$

単一化

Unify(E) が呼ばれたら, (E, \emptyset) から以下の書き換えを始め,
 (\emptyset, θ) の形に到達したらその θ を返す

途中で書き換えられなくなったら充足不可能 (型エラー相当)

$$(E \uplus \{\tau \stackrel{?}{=} \tau\}, \theta) \longrightarrow (E, \theta)$$

$$(E \uplus \{\tau \stackrel{?}{=} \alpha\}) \longrightarrow ([\tau / \alpha]E, \{\alpha \mapsto \tau\} \circ \theta)$$

(if $\alpha \notin \text{FTV}(\tau)$)

$$(E \uplus \{\tau_1 \rightarrow \tau_2 \stackrel{?}{=} \tau'_1 \rightarrow \tau'_2\}, \theta) \longrightarrow (E \cup \{\tau_1 \stackrel{?}{=} \tau'_1, \tau_2 \stackrel{?}{=} \tau'_2\}, \theta)$$

構成

- 背景
- おさらい：Let多相の型推論
- **[Ohori 1995]**
- [Osinski 2006]
- まとめ

機能面の概要（再掲）

- **SML#** の型システムの基礎づけ
- 双対である多相ヴァリアントもサポート（今回は扱わない）
- 一般的な入力に対して十分高速に型推論が動作
- ✱ レコードの拡張や接続はサポートしない
- ✓ 型推論可能
- ✓ コンパイルによる効率化が可能
- ✓ （やや主観ながら）推論される型が比較的わかりやすい

理論面の概要

- 型システム：
 - 有界量化による多相
 - 型の間の部分型関係 $<$: は陽には扱わず、
種 (kind) の形で表現する
- 型推論：
 - Let多相の型推論で使われる通常の単一化を
“種の制約解消と並行して行う単一化” へと変更して実現

項の定義の拡張

- 値 $v ::= c \mid \lambda x. e \mid \{\ell = v, \dots, \ell = v\}$
- 項 $e ::= v \mid x \mid e e \mid \text{let } x = e \text{ in } e$
 $\mid \{\ell = e, \dots, \ell = e\} \mid e \blacksquare \ell \mid e \text{ update } \ell e$

フィールド
取り出し

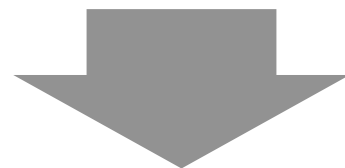
フィールド
更新

型付けのための考察

以下の関数にどんな型をつければよさそうか？

```
let send_missile = λaccount.  
  send_reply (account.screen_name) “🚀”
```

引数 *account* は, *screen_name* のフィールドが
文字列値として使われているのみ



**{screen_name= (文字列値), ...} の形の値は
全部受け取ってよい, これを反映した型をつける**

型と種：2階の体系

● 単相型 $\tau ::= b \mid \alpha \mid \tau \rightarrow \tau \mid \{l : \tau, \dots, l : \tau\}$

● 多相型 $\sigma ::= \tau \mid \forall \alpha :: \kappa. \sigma$

種による“制約”
つきの量化

● 種 $\kappa ::= U \mid \{l : \tau, \dots, l : \tau\}$

$\forall \alpha :: U. \sigma$ が従来の $\forall \alpha. \sigma$ に相当（つまり“制約”なし）

レコード種 $\{l_1 : \tau_1, \dots, l_n : \tau_n\}$ は直観的には

「少なくとも各 $l_i : \tau_i$ は全部含むレコード型全体」を指す

レコード種の直観

例

$$\{\text{foo} : \text{int}\} :: \{\{\text{foo} : \text{int}\}\}$$
$$\{\text{foo} : \text{int}, \text{bar} : \text{bool}\} :: \{\{\text{foo} : \text{int}\}\}$$
$$\{\text{foo} : \text{int}, \text{bar} : \text{bool}, \text{baz} : \text{string}\} :: \{\{\text{foo} : \text{int}\}\}$$

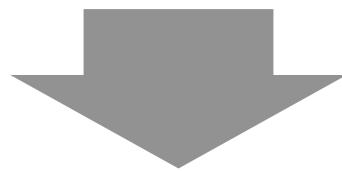
例

$$\{\text{foo} : \text{int}, \text{bar} : \text{bool}\} :: \{\{\text{foo} : \text{int}, \text{bar} : \text{bool}\}\}$$
$$\{\text{foo} : \text{int}, \text{bar} : \text{bool}, \\ \text{baz} : \text{string}\} :: \{\{\text{foo} : \text{int}, \text{bar} : \text{bool}\}\}$$

型つけの例

```
let send_missile =  $\lambda$ account.  
  send_reply (account.screen_name) “🚀”
```

引数 *account* の型は $\tau :: \{\text{screen_name} : \text{string}\}$
なる任意の τ でよい



send_missile につける型は

$$\forall \alpha :: \{\text{screen_name} : \text{string}\}. \alpha \rightarrow \text{unit}$$

量化に関する若干の注意

通常のLet多相と違い, 量化の順番も意味を持つ

```
let get_name =  $\lambda x. x.\text{name}$ 
```

get_name につく型は

$\forall \beta :: U. \forall \alpha :: \{\text{name} : \beta\}. \alpha \rightarrow \beta$ であって

$\forall \alpha :: \{\text{name} : \beta\}. \forall \beta :: U. \alpha \rightarrow \beta$ ではない

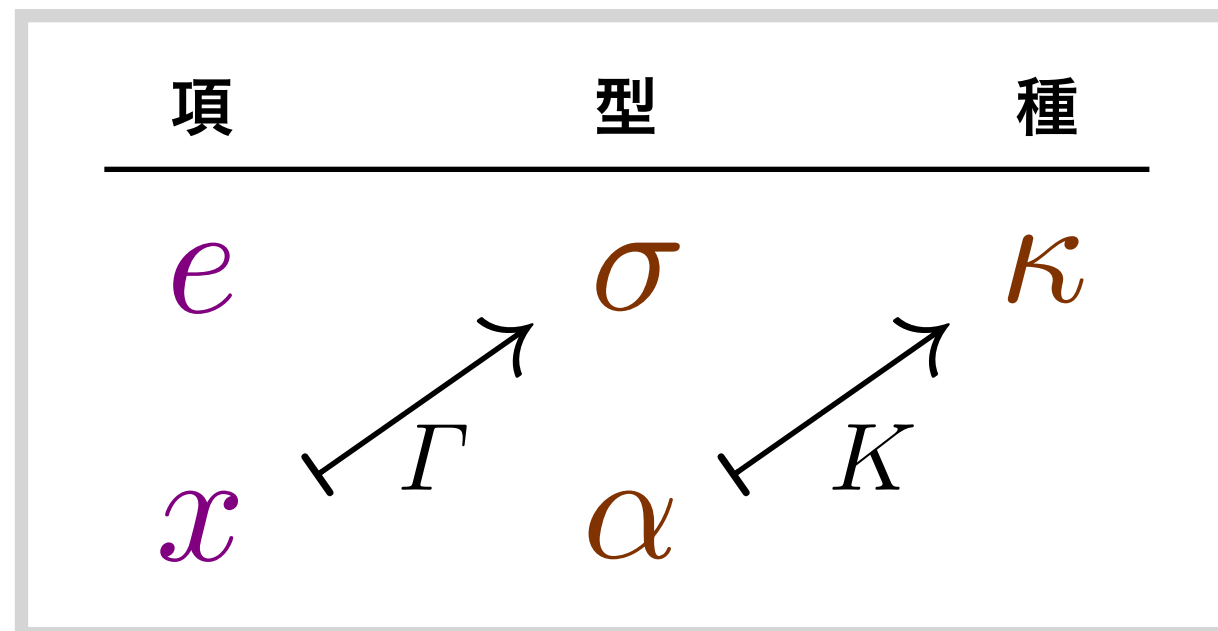
束縛されて
いない出現

型付け規則のための種環境の導入

種環境 K : 型変数に種を関連づける部分写像

$$K : \alpha \longmapsto \kappa$$

cf. 型環境 Γ : 変数に型を関連づける部分写像



型付け規則（抜粋1）

$$K \mid \Gamma \vdash e : \tau$$

“種環境 K の下で, τ は σ のインスタンスである”
(形式化は複雑なので割愛)

$$\frac{(\sigma \equiv \Gamma(x), \quad K \vdash \sigma \geq \tau)}{K \mid \Gamma \vdash x : \tau}$$

$$\frac{K \mid \Gamma \vdash e_i : \tau_i \quad (\text{for each } i)}{K \mid \Gamma \vdash \{\ell_1 = e_1, \dots, \ell_n = e_n\} : \{\ell_1 : \tau_1, \dots, \ell_n : \tau_n\}}$$

型付け規則 (抜粋2)

$$K \mid \Gamma \vdash e : \tau$$

“種環境 K の下で
 τ に種 $\{\ell : \tau'\}$ がつく”

$$\frac{K \mid \Gamma \vdash e : \tau \quad K \vdash \tau :: \{\ell : \tau'\}}{K \mid \Gamma \vdash (e \blacksquare \ell) : \tau'}$$
$$\frac{K \mid \Gamma \vdash e : \tau \quad K \mid \Gamma \vdash e' : \tau' \quad K \vdash \tau :: \{\ell : \tau'\}}{K \mid \Gamma \vdash (e \text{ update } \ell e') : \tau}$$

型付け規則 (抜粋2)

$$K \mid \Gamma \vdash e : \tau$$

“種環境 K の下で
 τ に種 $\{\ell : \tau'\}$ がつく”

$$\frac{K \mid \Gamma \vdash e : \tau \quad K \vdash \tau :: \{\ell : \tau'\}}{K \mid \Gamma \vdash (e \blacksquare \ell) : \tau'}$$

$$K \vdash \{R\} :: \{R'\} \\ \iff R \supseteq R'$$

$$K \vdash \alpha :: \{R'\} \\ \iff R \supseteq R' \\ (\{R\} \equiv K(\alpha))$$

$$\frac{K \mid \Gamma \vdash e : \tau \quad K \mid \Gamma \vdash e' : \tau' \quad K \vdash \tau :: \{\ell : \tau'\}}{K \mid \Gamma \vdash (e \text{ update } \ell e') : \tau}$$

型安全性

値呼び戦略では、計算が止まるときは必ず値になっている：

定理

$\emptyset \mid \emptyset \vdash e : \tau$ かつ $e \Downarrow e'$ ならば、 e' は τ 型の値である.

値呼び戦略で
簡約できなくなるまで
簡約し続ける

証明は論理関係 (logical relation) の手法による

型つけ規則から型推論アルゴリズムへ

この体系の型推論では、代入と共に種環境も降りてくる
(つまり種環境も型変数に対する制約の解消結果)

$$\frac{(\alpha \text{ fresh}) \quad K[\alpha \mapsto \mathbf{U}] \mid \Gamma[x \mapsto \alpha] \vdash e : \tau \downarrow K' \mid \theta}{K \mid \Gamma \vdash (\lambda x. e) : (\theta\alpha \rightarrow \tau) \downarrow K' \mid \theta}$$

$$\frac{\begin{array}{c} K \mid \Gamma \vdash e_1 : \tau_1 \downarrow K_1 \mid \theta_1 \quad (\beta_1, \beta \text{ fresh}) \\ (K', \theta') := \text{Unify}(K_1[\beta \mapsto \mathbf{U}][\beta_1 \mapsto \{\ell : \beta\}], \{\tau_1 \stackrel{?}{=} \beta_1\}) \end{array}}{K \mid \Gamma \vdash (e_1 \blacksquare \ell) : \theta'\beta \downarrow K' \mid \theta' \circ \theta_1}$$

種を加味した単一化（抜粋1）

$$(K, E \uplus \{\alpha_1 \stackrel{?}{=} \alpha_2\}, \theta) \longrightarrow (K', E', \theta')$$

where

$$\{\!\{R_1\}\!\} \equiv K(\alpha_1)$$

$$\{\!\{R_2\}\!\} \equiv K(\alpha_2)$$

両型変数の
現在の種

種を“合併”させる
(重複する部分は一方だけ)

$$K' := ([\alpha_2/\alpha_1]K)[\alpha_1 \mapsto [\alpha_2/\alpha_1]\{\!\{R_1 \cup R_2\}\!\}]$$

$$E' := [\alpha_2/\alpha_1]\left(E \cup \{R_1(\ell) \stackrel{?}{=} R_2(\ell) \mid \ell \in \text{dom } R_1 \cap \text{dom } R_2\}\right)$$

$$\theta' := ([\alpha_2/\alpha_1]\theta)[\alpha_1 \mapsto \alpha_2]$$

ラベルが重複する部分は
同一の型でないといけない

種を加味した単一化（抜粋2）

$$(K, E \uplus \{\alpha \stackrel{?}{=} \{R\}\}, \theta) \longrightarrow (K', E', \theta')$$

if $\text{dom } R' \subseteq \text{dom } R$ and $\alpha \notin \text{FTV}(\{R\})$

where

$$\{R'\} \equiv K(\alpha)$$

型変数の
現在の種

$$K' := [\{R\}/\alpha]K$$

$$E' := [\{R\}/\alpha] \left(E \cup \{R'(\ell) \stackrel{?}{=} R(\ell) \mid \ell \in \text{dom } R'\} \right)$$

$$\theta' := ([\{R\}/\alpha]\theta)[\alpha \mapsto \{R\}]$$

ラベルが重複する部分は
同一の型でないといけない

構成

- 背景
- おさらい：Let多相の型推論
- [Ohori 1995]
- **[Osinski 2006]**
- まとめ

機能面の概要（再掲）

- 博士論文（同著者唯一の publication らしい？）
- ✓ 前述のすべての演算をサポート
- ✓ 型推論可能
- ✓ コンパイルによる効率化が可能
- ▲ （主観ながら）推論される型が複雑でわかりにくい？
- ▲ 実装がまだ無いようで、**実時間の動作に堪えるか不明**
（自分の手で実装して確かめるつもりでしたが間に合わず）

理論面の概要

- 質化型 (qualified type) を土台としている：

$$\forall X. P \Rightarrow \tau$$

型変数の
集合

型変数が
満たす
述語の集合

- レコードやラベル集合の排反性や等価性を述語とし、
単一化に似た処理で述語集合の制約を解消
- アルゴリズムの正当性の証明するために
レコードの**多重集合による意味論**を定義しているっぽい

構成

- 背景
- おさらい：Let多相の型推論
- [Ohori 1995]
- [Osinski 2006]
- まとめ

まとめ

- レコード計算の型つけ, たのしい!
- [Ohori 1995]
 - SML# の型システムの基礎
 - 接続をサポートしないかわりに効率的に型推論が可能, 加えて効率的コードへとコンパイルできる
 - レコード計算のうちでは比較的型がわかりやすい?
- [Osinski 2006]
 - 機能面では “万能” らしい (証明未読)
 - 実装しないことには動作時間や使用感がわからない

蛇足

- 実は諏訪が実装した  という言語にも [Ohori 1995] の型システムが取り込まれています

<https://github.com/gfngfn/Macrodown>