Deriving a Modular and Complete Type Inference for Hindley-Milner with Row-Polymorphic Records by using Expansion: Proof Appendix

Axel Simon [*]

Technische Universität München
Lehrstuhl für Informatik 2
Garching b. München, Germany
Axel.Simon@in.tum.de

Abstract

Type inference and program analysis both infer static properties about a program. Yet, they are constructed using very different techniques. We reconcile both approaches by deriving a type inference from a denotational semantics using abstract interpretation. The novelty in our approach is the use of results from the abstract interpretation literature on completeness, modularity, and expansion to derive an inference that is abstract-complete, a property akin to the inference of principal typings. The resulting algorithm is simple but as powerful as that of Milner-Mycroft, that is, it infers Hindley-Milner types while allowing for polymorphic recursion. Results on modularity allow us to infer many polymorphic recursive types at the same asymptotic cost as Damas' $\mathcal{W}$-algorithm. We extend this algorithm with a complete and modular inference of Didier's row-polymorphic record types, that is, a program is rejected if it contains a path in which a record field is accessed without being set earlier. The inference uses Boolean functions and is, to our knowledge, the first complete type inference that is path sensitive.

Categories and Subject Descriptors F.3.2 [*Semantics of Programming Languages*]: Program analysis; F.3.3 [*Studies of Program Constructs*]: Type structure; D.3.1 [*Formal Definitions and Theory*]: Semantics; D.3.3 [*Language Constructs and Features*]: Polymorphism

General Terms principal typings, expansion, complete analysis, condensing domains, modular analysis, row polymorphism

Keywords type inference, abstract interpretation

A. Proof Appendix

This section is concerned with establishing that $\alpha(c_1) = \alpha(c_2) \Rightarrow \alpha(f(c_1)) = \alpha(f(c_2))$ holds in our context. We will first establish that the condition holds for sets of type vectors (Cor. 1) and then assert the same for polymorphic types (Lemma 2).

[*] Supported by ...

LEMMA 1. *For all $\rho_1, \rho_2 \in \mathbb{X} \to \mathbb{U}_\perp$ with $\alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_1) = \alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_2)$ it follows that $\alpha_{\mathbb{M}}(\mathcal{S}[\![e]\!]\ \rho_1) = \alpha_{\mathbb{M}}(\mathcal{S}[\![e]\!]\ \rho_2)$ for all expressions $e \in \mathbb{E}$.*

Proof. By structural induction over $e \in \mathbb{E}$:

x**:** $\alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_1)(x) = \mathcal{S}[\![x]\!]\ \rho_1 = \rho_1(x) = \rho_2(x) = \alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_2)(x)$.

$\lambda x\,.\,e$**:** Let $u_i = \mathcal{S}[\![e]\!]\ (\rho_i[x \mapsto v])$ for some $v \in \mathbb{U}$. Then $\alpha_{\mathbb{M}}(u_1) = \alpha_{\mathbb{M}}(u_2) := t_u$ by induction hypothesis. Thus $\alpha_{\mathbb{M}}(\mathcal{S}[\![\lambda x\,.\,e]\!]\ \rho_i) = t_v \to t_u$ for $i = 1, 2$.

$e^1\ e^2$**:** Let $v^j_i = \mathcal{S}[\![e^j]\!]\ \rho_i$, then $\alpha_{\mathbb{M}}(v^j_1) = \alpha_{\mathbb{M}}(v^j_2)$ for $j = 1, 2$ by induction hypothesis. Let $t_i = \alpha_{\mathbb{M}}(\mathcal{S}[\![e^1\ e^2]\!]\ \rho_i)$ for $i = 1, 2$. Assume $v^2_i \neq \Omega$ and $v^1_i \underline{\in} \mathbb{F}$ since otherwise $t_1 = t_2 = \emptyset$. Thus, $t_i = \alpha_{\mathbb{M}}(\downarrow^{\mathbb{U}}_{\mathbb{F}}(v^1_i)v^2_i)$ and, with i.h., $t_1 = t_2$ follows.

let $x = e$ **in** e'**:** By induction hypothesis $\alpha_{\mathbb{M}}(v'_1) = \alpha_{\mathbb{M}}(v'_2)$ where $v'_i = \mathcal{S}[\![e]\!]\ \rho_i[x \mapsto v])$ with $v \in \mathbb{U}$ it follows that $\alpha_{\mathbb{M}}(\hat{v}_1) = \alpha_{\mathbb{M}}(\hat{v}_2)$ where $\hat{v}_i = \mathrm{lfp}^{\preceq}_{\perp_{\mathbb{U}}} \lambda v\,.\,\mathcal{S}[\![e]\!]\ (\rho_i[x \mapsto v])$. By i.h. $\alpha_{\mathbb{M}}(v_1) = \alpha_{\mathbb{M}}(v_2)$ where $v_i = \mathcal{S}[\![\mathbf{let}\ x = e\ \mathbf{in}\ e']\!]\ \rho_i$.

Analogous for constructors and **case**-expressions. ■

Adjust Lemma 1 so that the results of the semantic action is stored in the environment, that is, show $\alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_1[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_1]) = \alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_2[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_2])$ instead of $\alpha_{\mathbb{M}}(\mathcal{S}[\![e]\!]\ \rho_1) = \alpha_{\mathbb{M}}(\mathcal{S}[\![e]\!]\ \rho_2)$. Lifting this result to sets yields the following observation:

COROLLARY 1. *For all $\bar{\rho}_1, \bar{\rho}_2 \subseteq \mathbb{X} \to \mathbb{U}_\perp$ with $\alpha^{\mathcal{X}}_{\mathbb{M}}(\bar{\rho}_1) = \alpha^{\mathcal{X}}_{\mathbb{M}}(\bar{\rho}_2)$ it follows that $\alpha^{\mathcal{X}}_{\mathbb{M}}(\{\rho_1[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_1] \mid \rho_1 \in \bar{\rho}_1\}) = \alpha^{\mathcal{X}}_{\mathbb{M}}(\{\rho_2[\kappa \mapsto \mathcal{S}[\![e]\!]\rho_2] \mid \rho_2 \in \bar{\rho}_2\})$ for all expressions $e \in \mathbb{E}$.*

By applying $\overline{\mathbf{lca}}$ on each side of the equations we obtain Lemma 2:

LEMMA 2. *For all $\bar{\rho}_1, \bar{\rho}_2 \subseteq \mathbb{X} \to \mathbb{U}_\perp$ with $\overline{\mathbf{lca}}(\alpha^{\mathcal{X}}_{\mathbb{M}}(\bar{\rho}_1)) = \overline{\mathbf{lca}}(\alpha^{\mathcal{X}}_{\mathbb{M}}(\bar{\rho}_2))$, it follows that $\overline{\mathbf{lca}}(\alpha^{\mathcal{X}}_{\mathbb{M}}(\{\rho_1[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_1] \mid \rho_1 \in \bar{\rho}_1\})) = \overline{\mathbf{lca}}(\alpha^{\mathcal{X}}_{\mathbb{M}}(\{\rho_2[\kappa \mapsto \mathcal{S}[\![e]\!]\rho_2] \mid \rho_2 \in \bar{\rho}_2\}))$ for all $e \in \mathbb{E}$.*

Proof. For the sake of a contradiction, let $t_i = \overline{\mathbf{lca}}(\alpha^{\mathcal{X}}_{\mathbb{M}}(\{\rho_i[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_i] \mid \rho_i \in \bar{\rho}_i\}))\kappa$ and $t_1 \neq t_2$. Then there exists $\rho^i_{\mathbb{M}} \in \mathscr{P}(\mathbb{X} \cup \{\kappa\} \to \mathbb{M})$ such that $\rho^i_{\mathbb{M}} \in \alpha^{\mathcal{X}}_{\mathbb{M}1}(\rho_i[\kappa \mapsto \mathcal{S}[\![e]\!]\ \rho_i])$ and $\rho^1_{\mathbb{M}}(\kappa) \neq \rho^2_{\mathbb{M}}(\kappa)$. However, this is a contradiction to Lem. 1. ■