An Abstract HOL-Kernel

Steven Obua *

Technische Universität München D-85748 Garching, Boltzmannstr. 3, Germany

An abstract HOL-kernel knows the primitive datatypes that are described in Figure 1. In Figure 2 we define the free *types* and *terms* of the kernel inductively

```
x \in \mathtt{nat} x is a non-negative integer x \in \mathtt{string} x \in \mathtt{id} x is a string x \in \mathtt{tag} x \in \mathtt{tag} x \in \mathtt{string} x \in \mathtt{tag} x \in \mathtt{string} x \in \mathtt{string}
```

Fig. 1. The primitive datatypes of an abstract HOL kernel

as sets type and term, respectively. The set of variables var and the set of constants const are subsets of term; the set of type variables tvar is a subset of type. Equality on elements of type and term is defined as usual by comparing the constructors and, if necessary, their arguments. Therefore $f \cdot g$ and $\lambda x.b$ are obviously not equal; and we have $\lambda x.b = \lambda x'.b'$ iff both x = x' and b = b' hold.

| $\alpha \in \mathtt{string}$ | $lpha\in {\sf tva}$ | r c | $\in \mathtt{id}$ | $T_1 \in {	t ty}$ | rpe · · · | $T_n \in {	t type}$ |
|--|---------------------------|--|---|--|---|--|
| $\overline{{\tt TVar}\ \alpha \in {\tt tvar}}$ | $\alpha \in \mathtt{typ}$ | $\alpha \in \mathtt{type}$ | | $\texttt{TApp}(c,T_1,\ldots,T_n)\in \texttt{type}$ | | |
| | | $\frac{x \in \mathtt{var}}{x \in \mathtt{term}}$ | _ | | $T \in \mathtt{type} \ T) \in \mathtt{const}$ | $\frac{c \in \mathtt{const}}{c \in \mathtt{term}}$ |
| $f \in \mathtt{term} \qquad g \in \mathtt{term}$ | | term | | $x \in \mathtt{var}$ | $b \in \mathtt{term}$ | m |
| $\overline{f\cdot g\in \mathtt{term}}$ | | | $\overline{\qquad \lambdax.b\in\mathtt{term}\qquad }$ | | _ | |

 ${\bf Fig.\,2.}$ Free types and terms

A kernel state S consists of

– the finite partial map $\operatorname{arity}_S \subseteq \operatorname{id} \times \operatorname{nat}$

 $^{^\}star$ Supported by the Ph.D. program "Logik in der Informatik" of the "Deutsche Forschungsgemeinschaft."

- and the finite partial map $ctype_S \subseteq id \times type$.

An HOL-kernel comes with certain reserved names for type and term constants. These names are elements of id and fixed for a given concrete kernel; but they are allowed to vary among different concrete kernels. We denote the two reserved type names by fun and bool. For each kernel state S we demand $arity_S(fun) = 2$ and $arity_S(bool) = 0$. The type of a term t is denoted $\tau(t)$ and defined as a partial map from term to type (fig. 3). In the rules for τ and in what follows we employ

$$\tau(\operatorname{Var}(x,T)) = T \qquad \qquad \tau(\operatorname{Const}(c,T)) = T$$

$$\frac{\tau(f) = A \to B \qquad \tau(g) = A}{\tau(f \cdot g) = B} \qquad \qquad \tau(\lambda \, x. \, b) = \tau(x) \to \tau(b)$$

Fig. 3. Type $\tau(t)$ of a term t

the abbreviations $A \to B := \operatorname{Tapp}(fun,A,B)$ and bool $:= \operatorname{Tapp}(bool)$. Note that as a consequence of the definition of τ , τ is not smart enough to perform matching in order to become less partial; for instance if $\tau(f) = \operatorname{TVar} \alpha \to \operatorname{bool}$ and $\tau(g) = \operatorname{bool}$, then $\tau(f \cdot g)$ is not defined.

For a given kernel state S, certain types and terms make sense and some don't. We call those which do make sense well-formed with respect to S and gather them into the sets \mathtt{type}_S and \mathtt{term}_S as defined in Figure 4. Special subsets

$$\frac{\alpha \in \mathsf{tvar}}{\alpha \in \mathsf{type}_S} \; (\mathsf{TVAR}) \qquad \frac{x \in \mathsf{string} \quad T \in \mathsf{type}_S}{\mathsf{Var} \, (x,T) \in \mathsf{term}_S} \; (\mathsf{VAR})$$

$$\frac{c \in \mathsf{id} \quad T_1 \in \mathsf{type}_S \quad \cdots \quad T_n \in \mathsf{type}_S \quad \mathit{arity}_S(c) = n}{\mathsf{TApp} \, (c,T_1,\ldots,T_n) \in \mathsf{type}_S} \; (\mathsf{TAPP})$$

$$\frac{c \in \mathsf{id} \quad T \in \mathsf{type}_S \quad \mathit{defined}_S(c,T') \quad T = \sigma T'}{\mathsf{Const} \, (c,T) \in \mathsf{term}_S} \; (\mathsf{Const})$$

$$\frac{f \in \mathsf{term}_S \quad g \in \mathsf{term}_S \quad \tau(f) = A \to B \quad \tau(g) = A}{f \cdot g \in \mathsf{term}_S} \; (\mathsf{APP})$$

$$\frac{\lambda \, x. \, b \in \mathsf{term} \quad \tau(x) \in \mathsf{type}_S \quad b \in \mathsf{term}_S}{\lambda \, x. \, b \in \mathsf{term}_S} \; (\mathsf{ABS})$$

Fig. 4. Well-formed types and terms with respect to a kernel state S

of term_S are the set of well-formed variables $\mathsf{var}_S := \mathsf{var} \cap \mathsf{term}_S$ and the set of well-formed $propositions\ \mathsf{prop}_S := \{ p \in \mathsf{term}_S \mid \tau(p) = \mathsf{bool} \}.$

In the Const-rule, σ is a type substitution, that is a finite partial map from tvar to type. There are also term substitutions, which are finite partial maps from var to term. For explicitly writing down a substitution σ we use the notation

$$\sigma = \langle redex_1 \mapsto residue_1, \dots, redex_n \mapsto residue_n \rangle$$
.

For type substitutions we have $redex_i \in \mathsf{tvar}$, $residue_i \in \mathsf{type}$, similarly for term substitutions $redex_i \in \mathsf{var}$, $residue_i \in \mathsf{term}$. In any case we adopt the convention that $redex_i = redex_j$ iff i = j. We call the sets of substitutions that involve only well-formed redexes and residues tsubst_S and subst_S , respectively. We reserve the notation $\sigma_{\backslash x}$ for the substitution that arises from σ by removing that redex/residue pair $redex \mapsto residue$ for which redex = x holds, if such a pair exists, and otherwise leaving the substitution unchanged.

A type substitution is extended in the obvious way to a total function on types and terms (fig. 5). Extending a term substitution to a total function on

$$\sigma \, \alpha = \begin{cases} T & \text{if } \alpha \mapsto T \text{ is a redex/residue pair of } \sigma \\ \alpha & \text{if } \alpha \in \mathsf{tvar} \text{ and there is no such redex/residue pair} \end{cases}$$

$$\sigma \left(\mathsf{TApp} \left(c, T_1, \dots, T_n \right) \right) = \mathsf{TApp} \left(c, \sigma \, T_1, \dots, \sigma \, T_n \right) \qquad \sigma \left(\mathsf{Var} \left(\alpha, T \right) \right) = \mathsf{Var} \left(\alpha, \sigma \, T \right)$$

$$\sigma \left(\mathsf{Const} \left(c, T \right) \right) = \mathsf{Const} \left(c, \sigma \, T \right) \qquad \sigma \left(f \cdot g \right) = \left(\sigma \, f \right) \cdot \left(\sigma \, g \right) \qquad \sigma \left(\lambda \, x. \, b \right) = \lambda \, \sigma \, x. \, \sigma \, b$$

Fig. 5. Applying a type substitution σ

terms is more involved because we have to be aware of name clashes of variables in λ -abstractions. We need the notion of *free variables* of terms and substitutions as defined in Figure 6. Then Figure 7 shows how term substitutions are applied.

$$\frac{x \in \text{var}}{\text{frees}\,(x) = \{x\}} \qquad \frac{c \in \text{const}}{\text{frees}\,(c) = \emptyset} \qquad \text{frees}\,(f \cdot g) = \text{frees}\,(f,g)$$

$$\text{frees}\,(\lambda \, x.\, b) = \text{frees}\,(b) - x$$

$$\text{frees}\,(t_1, \dots, t_n) = \text{frees}\,(\{t_1, \dots, t_n\}) = \text{frees}\,(t_1) \cup \dots \cup \text{frees}\,(t_n)$$

$$\text{frees}\,(\langle \, x_1 \mapsto t_1, \dots, \, x_n \mapsto t_n \, \rangle) = \{x_1, \dots, x_n\} \cup \text{frees}\,(t_1, \dots, t_n)$$

Fig. 6. Free variables of terms and term substitutions

Obviously the result r of applying a substitution σ to a term t depends on the choices made for y in the λ -case. Therefore r is not a term, but rather a set of terms, the set of all possible outcomes. Of course we will forget about this fact and just pick one of the possible outcomes, but for the next definition only let us write $Outcomes(\sigma,t)$ for this set. Thus we say that two terms s and t are α -equivalent iff applying the empty substitution to both terms yields the same set of outcomes:

$$s \stackrel{\alpha}{=} t$$
 iff $Outcomes(\langle \rangle, s) = Outcomes(\langle \rangle, t)$.

The abstract HOL-kernel and all the concrete ones always operate modulo at

$$\frac{c \in \mathsf{const}}{\sigma \, c = c} \qquad \sigma \, x = \begin{cases} t & \text{if } x \mapsto t \text{ is a redex/residue pair of } \sigma \\ x & \text{if } x \in \mathsf{var} \text{ and there is no such redex/residue pair} \end{cases}$$

$$\sigma \, (f \cdot g) = (\sigma \, f) \cdot (\sigma \, g)$$

$$\sigma \, (\lambda \, x. \, b) = \lambda \, y. \, \sigma_{\backslash x} \, (\langle \, x \mapsto y \, \rangle \, b), \text{ where } y \in \mathsf{var}, \, \tau \, (x) = \tau \, (y) \text{ and } y \notin \mathsf{frees} \, (\sigma_{\backslash x}, b)$$

Fig. 7. Applying a term substitution σ

least α -equivalence on terms, therefore there is no harm in considering σt a term if t is a term. To add to your confusion, we do consider $\sigma \Gamma$ a set if $\Gamma = \{t_1, \ldots, t_n\}$ is a set of terms, defined in the obvious way by $\sigma \Gamma = \{\sigma t_1, \ldots, \sigma t_n\}$.

Now we are finally in a position to deal with the two most important notions of our kernel, theorems and proofs. A theorem is always considered with respect to a kernel state S, consisting of a set Γ of hypotheses together with a conclusion c, all of which are well-formed propositions with respect to S; and it is always accompanied by a proof P that tells us which rules were used in what order to wrest the theorem from the kernel. We write such a theorem/proof combination

$$P \leadsto_S \Gamma \vdash c$$
 or, if Γ is the empty set $P \leadsto_S c$.

So the above notation implies:

- $-\Gamma\subseteq\operatorname{prop}_S,\ c\in\operatorname{prop}_S;$ also all the term and types mentioned in P are well-formed relative to S;
- S is the kernel state that is obtained as a side-effect of replaying the proof P starting in the *initial state* of the HOL-kernel;
- the theorem $\Gamma \vdash c$ can be deduced using the rules of the kernel, which again can be checked by replaying the proof P.

TODO:

 Explain better: The initial state of the HOL-kernel is not clearly defined; it can vary from system to system, and therefore we also do not fix it.

- define tfrees (t) define TyDef, TyIntro, side-effects of them and of Def side-effects that do not

$$\frac{u \in \mathsf{term}_S}{\mathsf{REFL}(u) \leadsto_S u \doteq u} \qquad \frac{a \in \mathsf{prop}_S}{\mathsf{ASSUME}(a) \leadsto_S \{a\} \vdash a}$$

$$\frac{P \leadsto_S \Gamma \vdash u \doteq (\lambda x.b) \cdot t}{\mathsf{BETA}(P) \leadsto_S \Gamma \vdash u \doteq (x \mapsto t) b} \qquad \frac{P \leadsto_S \Gamma \vdash u \doteq \lambda x.f \cdot x \quad x \notin \mathsf{frees}(f)}{\mathsf{ETA}(P) \leadsto_S \Gamma \vdash u \doteq f}$$

$$\frac{P \leadsto_S \Gamma \vdash a \doteq b \quad Q \leadsto_S \Gamma' \vdash b \doteq c}{\mathsf{TRANS}(P,Q) \leadsto_S \Gamma \cup \Gamma' \vdash b} \qquad \frac{P \leadsto_S \Gamma \vdash a \doteq b}{\mathsf{SYM}(P) \leadsto_S \Gamma \vdash b \doteq a}$$

$$\frac{P \leadsto_S \Gamma \vdash a \to b \quad Q \leadsto_S \Gamma' \vdash a}{\mathsf{MP}(P,Q) \leadsto_S \Gamma \cup \Gamma' \vdash b} \qquad \frac{P \leadsto_S \Gamma \vdash a \doteq b}{\mathsf{EQMP}(P,Q) \leadsto_S \Gamma \cup \Gamma' \vdash b}$$

$$\frac{P \leadsto_S \Gamma \vdash a \doteq b}{\mathsf{EQIMP}(P) \leadsto_S \Gamma \vdash a \to b} \qquad \frac{P \leadsto_S \Gamma \vdash a \doteq b}{\mathsf{IMPANTISYM}(P,Q) \leadsto_S \Gamma \cup \Gamma' \vdash b}$$

$$\frac{P \leadsto_S \Gamma \vdash a \quad Q \leadsto_S \Gamma' \vdash b}{\mathsf{EQIMP}(P,Q) \leadsto_S \Gamma \cup \Gamma' \vdash a \land b} \qquad \frac{P \leadsto_S \Gamma \vdash a \land b}{\mathsf{CONJI}(P) \leadsto_S \Gamma \vdash a \to b}$$

$$\frac{P \leadsto_S \Gamma \vdash a \quad \to b}{\mathsf{NOTI}(P) \leadsto_S \Gamma \vdash a \land b} \qquad \frac{P \leadsto_S \Gamma \vdash a \land b}{\mathsf{NOTI}(P) \leadsto_S \Gamma \vdash a \lor b} \qquad \frac{P \leadsto_S \Gamma \vdash a \land b}{\mathsf{NOTI}(P) \leadsto_S \Gamma \vdash a \lor b}$$

$$\frac{P \leadsto_S \Gamma \vdash a \quad b \in \mathsf{prop}_S}{\mathsf{DISJ}(P,b) \leadsto_S \Gamma \vdash a \lor b} \qquad \frac{P \leadsto_S \Gamma \vdash b \quad a \in \mathsf{prop}_S}{\mathsf{DISJ}(P,b) \leadsto_S \Gamma \vdash a \lor b}$$

$$\frac{P \leadsto_S \Gamma \vdash a \lor b}{\mathsf{DISJ}(\mathsf{CASES}(P,Q,R) \leadsto_S \Gamma \cup (\Gamma' \vdash a) \cup (\Gamma'' \vdash a) \vdash c}$$

$$\frac{P \leadsto_S \Gamma \vdash c \quad h \in \mathsf{prop}_S}{\mathsf{DIOL}(t) \leadsto_S (t \doteq 0) \lor (t \doteq 1)} \qquad \frac{P \leadsto_S \Gamma \vdash b \quad a \in \mathsf{prop}_S}{\mathsf{INJI}(P,f) \leadsto_S \Gamma \vdash inj f}$$

$$\frac{P \leadsto_S \Gamma \vdash f \lor_S f \vdash g \lor_S f \vdash f \lor_S f \vdash g \lor_S f \vdash g \lor_S f \vdash g \lor_S f \vdash g \lor_S f \vdash_S f \vdash_S f \lor_S f$$

Fig. 8. Rules without binders

Fig. 9. Higher-order rules

$$\begin{split} \frac{P \leadsto_S \Gamma \vdash c \qquad \sigma = \langle \, \alpha_1 \mapsto T_1, \dots, \, \alpha_n \mapsto T_n \, \rangle \in \mathtt{tsubst}_S}{\mathtt{INSTTYPE} \, (P, \alpha_1, T_1, \dots, \alpha_n, T_n) \leadsto_S \sigma \, \Gamma \vdash \sigma \, c} \\ \\ \frac{P \leadsto_S \Gamma \vdash c \qquad \sigma = \langle \, x_1 \mapsto t_1, \dots, \, x_n \mapsto t_n \, \rangle \in \mathtt{subst}_S}{\mathtt{INST} \, (P, x_1, t_1, \dots, x_n, t_n) \leadsto_S \sigma \, \Gamma \vdash \sigma \, c} \\ \\ \frac{P_1 \leadsto_S \Gamma_1 \vdash l_1 \stackrel{.}{=} r_1 \quad \dots \quad P_n \leadsto_S \Gamma_n \vdash l_n \stackrel{.}{=} r_n \quad t \in \mathtt{term}_S}{\forall \, \sigma_\gamma = \langle \, x_1 \mapsto \gamma_1, \dots, \, x_n \mapsto \gamma_n \, \rangle \in \mathtt{subst}_S} \quad P \leadsto_S \Gamma \vdash \sigma_l \, t} \\ \\ \frac{\forall \, \sigma_\gamma = \langle \, x_1 \mapsto \gamma_1, \dots, \, x_n \mapsto \gamma_n \, \rangle \in \mathtt{subst}_S}{\mathtt{SUBST} \, (P_1, \dots, P_n, \, \lambda \, x_1, \dots \lambda \, x_n, t, P) \leadsto_S \Gamma \cup \Gamma_1 \cup \dots \cup \Gamma_n \vdash \sigma_r \, t} \end{split}$$

Fig. 10. Substitution rules

$$\frac{s \in \mathtt{string} \quad c \in \mathtt{prop}_S}{\mathtt{AXIOM}\,(s,c) \leadsto_S c} \qquad \frac{s \in \mathtt{tag} \qquad h_1, \dots, h_n, c \in \mathtt{prop}_S}{\mathtt{ORACLE}\,(s,h_1,\dots,h_n,c) \leadsto_S \{h_1,\dots,h_n\} \vdash c}$$

$$\frac{s \in \mathtt{tag} \qquad h_1,\dots,h_n \in \mathtt{prop}_S \qquad c \in \mathtt{prop}_S}{T_1,\dots,T_v \in \mathtt{type}_S \qquad P_1 \leadsto_S \Gamma_1 \vdash c_1 \qquad \dots \qquad P_w \leadsto_S \Gamma_w \vdash c_w}$$

$$\underline{t_1,\dots,t_u \in \mathtt{term}_S \qquad T_1,\dots,T_v \in \mathtt{type}_S \qquad P_1 \leadsto_S \Gamma_1 \vdash c_1 \qquad \dots \qquad P_w \leadsto_S \Gamma_w \vdash c_w}}$$

$$\underline{\mathtt{METHOD}\,(s,h_1,\dots,h_n,c,t_1,\dots,t_u,T_1,\dots,T_v,P_1,\dots,P_w) \leadsto_S \{h_1,\dots,h_n\} \vdash c}$$

Fig. 11. Unsafe rules

$$\frac{P \leadsto_S \{h_1, \dots, h_n\} \vdash c \qquad c', h'_1, \dots, h'_n \in \mathsf{term}_S \qquad c \cong c' \quad \forall h_i \cong h'_i}{P \leadsto_S \{h'_1, \dots, h'_n\} \vdash c'} \ (\mathsf{Modulo})$$

Fig. 12. Rule for proving modulo \cong

$$\begin{array}{ccc} c \in \mathtt{id} & c \notin \mathrm{dom}\,(ctype_S) & t \in \mathtt{term}_S \\ & \underline{ & \mathrm{frees}\,(t) = \emptyset & \mathrm{tfrees}\,(t) = \mathrm{tfrees}\,(\tau(t)) \\ \hline & \underline{ & \mathrm{DEF}\,(c,t) \leadsto_{S'} \mathtt{Const}\,(c,\tau(t)) \doteq t } \end{array}) \end{array}$$

Fig. 13. State-changing rules (TO DO: document TyDef, TyIntro)