Imperative HOL – a leightweight framework for imperative data structures in Isabelle/HOL

August 27, 2014

Imperative HOL is a leightweight framework for reasoning about imperative data structures in Isabelle/HOL [2]. Its basic ideas are described in [1]. However their concrete realisation has changed since, due to both extensions and refinements. Therefore this overview wants to present the framework “as it is” by now. It focusses on the user-view, less on matters of construction. For details study of the theory sources is encouraged.

1 A polymorphic heap inside a monad

Heaps (heap) can be populated by values of class heap; HOL’s default types are already instantiated to class heap. Class heap is a subclass of countable; see theory Countable for ways to instantiate types as countable.

The heap is wrapped up in a monad 'a Heap by means of the following specification:

datatype 'a Heap = Heap

Unwrapping of this monad type happens through

execute :: 'a Heap ⇒ heap ⇒ ('a × heap) option
execute (Heap.Heap f) = f

This allows for equational reasoning about monadic expressions; the fact collection execute-simps contains appropriate rewrites for all fundamental operations.

Primitive fine-granular control over heaps is available through rule Heap-cases:

(∃x h'. execute f h = Some (x, h') ⇒ P) ⇒ (execute f h = None ⇒ P) ⇒ P
Monadic expression involve the usual combinators:

\[
\begin{align*}
\text{return} & :: 'a \Rightarrow 'a \text{Heap} \\
\text{op } \gg= & :: 'a \text{Heap} \Rightarrow ('a \Rightarrow 'b \text{Heap}) \Rightarrow 'b \text{Heap} \\
\text{raise} & :: 'a \text{list} \Rightarrow 'a \text{Heap}
\end{align*}
\]

This is also associated with nice monad do-syntax. The string argument to \text{raise} is just a codified comment.

Among a couple of generic combinators the following is helpful for establishing invariants:

\[
\begin{align*}
\text{assert} & :: ('a \Rightarrow \text{bool}) \Rightarrow 'a \Rightarrow 'a \text{Heap} \\
\text{assert } P \ x & = (\text{if } P \ x \text{ then return } x \text{ else raise } "\text{assert}")
\end{align*}
\]

\section{Relational reasoning about Heap expressions}

To establish correctness of imperative programs, predicate

\[
\text{effect} :: 'a \text{Heap} \Rightarrow \text{heap} \Rightarrow \text{heap} \Rightarrow 'a \Rightarrow \text{bool}
\]

provides a simple relational calculus. Primitive rules are \text{effectI} and \text{effectE}, rules appropriate for reasoning about imperative operations are available in the \text{effect-intros} and \text{effect-elims} fact collections.

Often non-failure of imperative computations does not depend on the heap at all; reasoning then can be easier using predicate

\[
\text{success} :: 'a \text{Heap} \Rightarrow \text{heap} \Rightarrow \text{bool}
\]

Introduction rules for \text{success} are available in the \text{success-intro} fact collection.

\text{execute}, \text{effect}, \text{success} and \text{op } \gg= \text{ are related by rules execute-bind-success, success-bind-executeI, success-bind-effectI, effect-bindI, effect-bindE and execute-bind-eq-SomeI.}

\section{Monadic data structures}

The operations for monadic data structures (arrays and references) come in two flavours:

\begin{itemize}
\item Operations on the bare heap; their number is kept minimal to facilitate proving.
\end{itemize}
- Operations on the heap wrapped up in a monad; these are designed for executing.

Provided proof rules are such that they reduce monad operations to operations on bare heaps.

Note that HOL equality coincides with reference equality and may be used as primitive executable operation.

### 3.1 Arrays

Heap operations:

\[
\begin{align*}
\text{Array.alloc} & : \text{'a list} \Rightarrow \text{heap} \Rightarrow \text{'a array } \times \text{heap} \\
\text{Array.present} & : \text{heap} \Rightarrow \text{'a array } \Rightarrow \text{bool} \\
\text{Array.get} & : \text{heap} \Rightarrow \text{'a array } \Rightarrow \text{'a list} \\
\text{Array.set} & : \text{'a array } \Rightarrow \text{'a list } \Rightarrow \text{heap } \Rightarrow \text{heap} \\
\text{Array.length} & : \text{heap} \Rightarrow \text{'a array } \Rightarrow \text{nat} \\
\text{Array.update} & : \text{'a array } \Rightarrow \text{nat } \Rightarrow \text{'a } \Rightarrow \text{heap } \Rightarrow \text{heap} \\
\text{op }=!= & : \text{'a array } \Rightarrow \text{'b array } \Rightarrow \text{bool}
\end{align*}
\]

Monad operations:

\[
\begin{align*}
\text{Array.new} & : \text{nat } \Rightarrow \text{'a } \Rightarrow \text{'a array} \text{ Heap} \\
\text{Array.of-list} & : \text{'a list } \Rightarrow \text{'a array} \text{ Heap} \\
\text{Array.make} & : \text{nat } \Rightarrow (\text{nat } \Rightarrow \text{'a}) \Rightarrow \text{'a array} \text{ Heap} \\
\text{Array.len} & : \text{'a array } \Rightarrow \text{nat} \text{ Heap} \\
\text{Array.nth} & : \text{'a array } \Rightarrow \text{nat } \Rightarrow \text{'a} \text{ Heap} \\
\text{Array.upd} & : \text{nat } \Rightarrow \text{'a } \Rightarrow \text{'a array } \Rightarrow \text{'a array} \text{ Heap} \\
\text{Array.map-entry} & : \text{nat } \Rightarrow (\text{'a } \Rightarrow \text{'a}) \Rightarrow \text{'a array } \Rightarrow \text{'a array} \text{ Heap} \\
\text{Array.swap} & : \text{nat } \Rightarrow \text{'a } \Rightarrow \text{'a array } \Rightarrow \text{'a} \text{ Heap} \\
\text{Array.freeze} & : \text{'a array } \Rightarrow \text{'a list} \text{ Heap}
\end{align*}
\]

### 3.2 References

Heap operations:

\[
\begin{align*}
\text{Ref.alloc} & : \text{'a } \Rightarrow \text{heap } \Rightarrow \text{'a ref } \times \text{heap} \\
\text{Ref.present} & : \text{heap } \Rightarrow \text{'a ref } \Rightarrow \text{bool} \\
\text{Ref.get} & : \text{heap } \Rightarrow \text{'a ref } \Rightarrow \text{'a} \\
\text{Ref.set} & : \text{'a ref } \Rightarrow \text{'a } \Rightarrow \text{heap } \Rightarrow \text{heap} \\
\text{op }=!= & : \text{'a ref } \Rightarrow \text{'b ref } \Rightarrow \text{bool}
\end{align*}
\]

Monad operations:
Imperative HOL sets up the code generator in a way that imperative operations are mapped to suitable counterparts in the target language. For Haskell, a suitable ST monad is used; for SML, Ocaml and Scala unit values ensure that the evaluation order is the same as you would expect from the original monadic expressions. These units may look cumbersome; the target language variants SML-imp, Ocaml-imp and Scala-imp make some effort to optimize some of them away.

5 Some hints for using the framework

Of course a framework itself does not by itself indicate how to make best use of it. Here some hints drawn from prior experiences with Imperative HOL:

- Proofs on bare heaps should be strictly separated from those for monadic expressions. The first capture the essence, while the latter just describe a certain wrapping-up.

- A good methodology is to gradually improve an imperative program from a functional one. In the extreme case this means that an original functional program is decomposed into suitable operations with exactly one corresponding imperative operation. Having shown suitable correspondence lemmas between those, the correctness prove of the whole imperative program simply consists of composing those.

- Whether one should prefer equational reasoning (fact collection execute-simps or relational reasoning (fact collections effect-intros and effect-elims) depends on the problems to solve. For complex expressions or expressions involving binders, the relation style usually is superior but requires more proof text.

- Note that you can extend the fact collections of Imperative HOL yourself whenever appropriate.
References
