Composable Packages for Higher Order Logic Theories

Joe Hurd

Galois, Inc.
joe@galois.com

VERIFY Workshop
Wednesday 21 July 2010

Talk Plan

galois

Motivation

- Interactive theorem proving is growing up.
  - The FlySpeck project is driving the HOL Light theorem prover towards a formal proof of the Kepler sphere-packing conjecture.
  - The seL4 project recently completed a 20 man-year verification of an operating system kernel in the Isabelle theorem prover.
- There is a need for theory engineering techniques to support these major verification efforts.
  - Theory engineering is to proving as software engineering is to programming.
  - *"Proving in the large."*

The OpenTheory Project

- The goal of the OpenTheory project is to transfer the benefits of package management to logical theories.[1]
- The initial case study for the project is Church's simple theory of types, extended with Hindley-Milner style type variables.
  - The logic implemented by HOL4, HOL Light and ProofPower.
- By focusing on a concrete case study we aim to investigate the issues surrounding:
  - Designing theory languages portable across theorem prover implementations.
  - Discovering design techniques for reusable theories.
  - Uploading, installing and upgrading theory packages from online repositories.
  - Building a standard theory library.

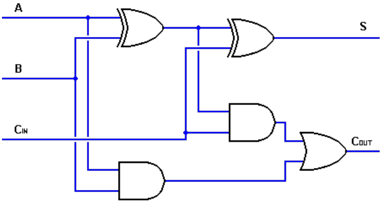
[1] OpenTheory was started in 2004 with Rob Arthan.

Theory Definition

- A theory $\Gamma \rhd \Delta$ of higher order logic consists of:
  1. A set Γ of assumption sequents.
  2. A set Δ of theorem sequents.
  3. A formal proof that the theorems in Δ logically derive from the assumptions in Γ.
- Theories (including their proofs) can be directly represented as OpenTheory article files.
  - A format designed to simplify theory import and export for theorem prover implementations.
- This talk will present a language for building up from article files to theory packages.
  - We'll see toy case studies that demonstrate the concepts, but the true test will be whether it scales up—watch this space!

Connecting Theories

- Note that both the input assumptions and output theorems of a theory are sequent sets.
- We can therefore connect the output theorems of one theory to satisfy the input assumptions of another:



- In this illustration, some theories have been connected together to produce the compound theory

$$A \cup B \cup C_{\mathrm{IN}} \triangleright S \cup C_{\mathrm{OUT}} \ .$$

Theory Interpretations

- A theory $\Gamma \rhd \Delta$ can be instantiated in any context where the assumptions Γ hold. This is called theory interpretation.
- Example: The theory

$$\{\vdash \mathrm{id} = \lambda x.\ x\} \rhd \{\vdash \forall x.\ \mathrm{id}\ x = x\}$$

  can be applied in any context with a constant id having the assumed property.
- Constants and type operators can be consistently renamed

$$(\Gamma \rhd \Delta)\sigma \;=\; \Gamma\sigma \rhd \Delta\sigma$$

  allowing theories to be instantiated in even more contexts.

What Can Go Wrong?

- When connecting together theories, the connection graph must not contain any loops!
  - Theories are representations of proofs, which are directed *acyclic* graphs.
  - In this aspect proofs are more like combinational circuits than programs.
- A set of theorems must not have incompatible definitions for the same constant or type operator.
  - Example: The two theories

    $$\emptyset \rhd \{\vdash \mathsf{c} = 0\} \qquad \text{and} \qquad \emptyset \rhd \{\vdash \mathsf{c} = 1\}$$

    are individually fine, but must never be imported into the same context.

galois

A Language for Theories

- The following theory language allows article files and theory packages to be combined into a new theory:

theory	←	`article "`*filename*`";`
	|	`{` *theory** `}`
	|	`local` *theory* `in` *theory*
	|	`interpret {` *interpretation** `} in` *theory*
	|	`import` *package-instance*`;`

- Incompatible definition clashes are prevented by:
  - Limiting the scope of contexts using the `local` construct.
  - Renaming constant and type operators using `interpret` blocks.

galois

Theory Package Example

Theory Package (hol-light-unit-def-2009.8.24)

```
name: hol-light-unit-def
version: 2009.8.24
description: HOL Light definition of the unit type.

theory { article "hol-light-unit-def.art"; }
```

galois

Theory Package Example

Theory Package Summary (hol-light-unit-def-2009.8.24)

```
input-types: -> bool
input-consts: ! /\ = ? T select
assumed:
  |- T
  {.} |- (!) P
  {.} |- (?) P
  {..} |- p /\ q
  |- t = (t = T)
  |- (?) = \P. P ((select) P)
defined-types: unit
defined-consts: one one_ABS one_REP
thms:
  |- ?b. b
  |- one = select x. T
  |- (!a. one_ABS (one_REP a) = a) /\
     !r. r = (one_REP (one_ABS r) = r)
```

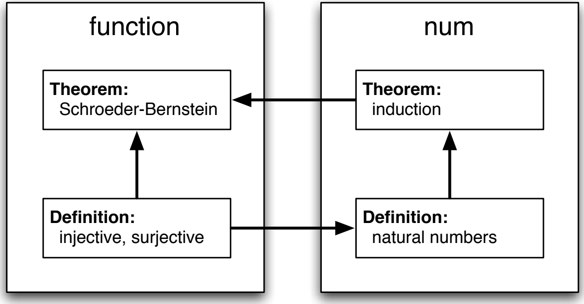
Theory Package Design

- Well-designed theory packages have:
  1. A clear topic.
     - Example: Trigonometric functions.
  2. A simple set of assumptions.
     - Satisfied by well-designed packages.
  3. A carefully chosen set of theorems.
     - No junk.
     - A minimal interface if the package makes definitions.
  4. No axioms.
     - No assumptions about defined constants/type operators.
- Theory Engineering Challenge: Construct a standard library of well-designed theory packages, available to all the HOL theorem prover implementations.

galois

Theory Dependencies

- **Problem:** Complex theory dependencies can result in cycles in the package dependency graph.



- **Solution:** Permit compilation theory packages which contain previously loaded theory packages.

galois

Theory Package Instances

- An imported *package-instance* refers to a required theory package, specified as a *package-instance-spec*:

package-instance-spec ← `require` *package-instance* {
  `import:` *package-instance**
  `interpret:` *interpretation**
  `package:` *package-name*
}

- A list of *package-instance-spec*s specify a connection graph between theory packages.
- Each *package-instance-spec* may only import earlier *package-instance-spec*s, to ensure the absence of loops.

Theory Packages

- We can now define the grammar for theory packages:

$$\begin{array}{rcl} \textit{package} & \leftarrow & \textit{tag}^* \\ & & \textit{package-instance-spec}^* \\ & & \texttt{theory} \ \{ \ \textit{theory} \ \} \end{array}$$

- Tags are package meta-data:

$$\textit{tag} \quad \leftarrow \quad \textit{name} \texttt{:} \ \textit{value}$$

Theory Package Example II

Theory Package (unit-def-1.0)

```
name: unit-def
version: 1.0
description: Definition of the unit type

require hol-light-aux {
  package: hol-light-aux-2009.8.24
}

require hol-light-unit-def {
  import: hol-light-aux
  package: hol-light-unit-def-2009.8.24
}

require hol-light-unit-alt {
  import: hol-light-aux
  import: hol-light-unit-def
  package: hol-light-unit-alt-2009.8.24
}

theory { import hol-light-unit-alt; }
```

Theory Package Example II

Theory Package Summary (unit-def-1.0)

```
input-types: -> bool
input-consts: ! /\ = ==> ? T select
assumed:
  |- !t. (\x. t x) = t
  |- T = ((\p. p) = \p. p)
  |- (!) = \P. P = \x. T
  |- (==>) = \p q. (p /\ q) = p
  |- !P x. P x ==> P ((select) P)
  |- (/\) = \p q. (\f. f p q) = \f. f T T
  |- (?) = \P. !q. (!x. P x ==> q) ==> q
defined-types: unit
defined-consts: one
thms:
  |- !v. v = one
```

Symbol Tables Considered Harmful

- To make it easy to reason about theory package instances, we would like package instantiation to be a pure function

$$\textit{package-instance-spec} \longrightarrow \Gamma \triangleright \Delta \ .$$

- Possible because the package management tool implements a purely functional logical kernel (an idea of Freek Wiedijk).
- Constants and type operators contain their definitions, instead of being inserted in a symbol table, so definitions are referentially transparent:

$$(\text{let } c = \text{define } \phi \text{ in } f\ c\ c) \ \equiv \ (f\ (\text{define } \phi)\ (\text{define } \phi))$$

galois

Efficient Sharing

- Referential transparency means there is no difference in functionality between instantiating a theory package multiple times in the same way or instantiating it once and reusing.
- However, there will likely be a big difference in performance (article files are measured in megabytes).
- Challenge: Detecting when two *package-instance-spec*s would result in the same theory.
- The logical kernel similarly aims to share subterms as much as possible, in computing free variables, substitutions, etc.

Summary

- This talk presented a language for combining and packaging theories.
- The next challenge: build the package management infrastructure for people to contribute to building a standard library of theories.
- The project web page:

  `http://gilith.com/research/opentheory`

galois

Package Instance Semantics

- The concrete syntax for *package-instance-spec* evaluates to the theory

$$\bigcup \Gamma_i \cup \left(\Gamma\sigma - \bigcup \Delta_i \right) \rhd \Delta\sigma$$

  where:

  - the imported *package-instance-spec*s evaluate to $\Gamma_i \rhd \Delta_i$;
  - the *interpretation* rules are the renaming σ; and
  - the *package-name* is the theory $\Gamma \rhd \Delta$.

Theory Semantics

- Here is how the concrete syntax for *theory* is evaluated in a context with theorems Φ and renaming σ:

$$
\begin{array}{rcl}
[\texttt{article "}[\Gamma \triangleright \Delta]\texttt{";}]_{\Phi,\sigma} & = & \Gamma\sigma - \Phi \triangleright \Delta\sigma \\
[\{\ []\ \}]_{\Phi,\sigma} & = & \emptyset \triangleright \emptyset \\
[\{\ \theta_1 :: \theta_2\ \}]_{\Phi,\sigma} & = & \text{let } \Gamma_1 \triangleright \Delta_1 = [\theta_1]_{\Phi,\sigma} \text{ in} \\
 & & \text{let } \Gamma_2 \triangleright \Delta_2 = [\{\ \theta_2\ \}]_{\Phi \cup \Delta_1,\sigma} \text{ in} \\
 & & \Gamma_1 \cup \Gamma_2 \triangleright \Delta_1 \cup \Delta_2 \\
[\texttt{local } \theta_1 \texttt{ in } \theta_2]_{\Phi,\sigma} & = & \text{let } \Gamma_1 \triangleright \Delta_1 = [\theta_1]_{\Phi,\sigma} \text{ in} \\
 & & \text{let } \Gamma_2 \triangleright \Delta_2 = [\theta_2]_{\Phi \cup \Delta_1,\sigma} \text{ in} \\
 & & \Gamma_1 \cup \Gamma_2 \triangleright \Delta_2 \\
[\texttt{interpret } \{\ \rho\ \} \texttt{ in } \theta]_{\Phi,\sigma} & = & [\theta]_{\Phi,\sigma \circ \rho} \\
[\texttt{import } [\Gamma \triangleright \Delta]\texttt{;}]_{\Phi,\sigma} & = & \Gamma \triangleright \Delta
\end{array}
$$

- Note that importing a *package-instance* ignores the theory context; its context is fixed by the *package-instance-spec*.