Interactive Theorem Proving (ITP) Course
Parts I - IV

Thomas Tuerk (tuerk@kth.se)

KTH

Academic Year 2016/17, Period 4

version 1b43e90 of Sun Apr 23 20:08:16 2017

Part I

Introduction

Motivation

- Complex systems almost certainly contain bugs.
- Critical systems (e. g. avionics) need to meet very high standards.
- It is infeasible in practice to achieve such high standards just by testing.
- Debugging via testing suffers from diminishing returns.

"Program testing can be used to show the presence of bugs, but never to show their absence!"
— Edsger W. Dijkstra

Famous Bugs

- Pentium FDIV bug (1994)
  (missing entry in lookup table, $475 million damage)
- Ariane V explosion (1996)
  (integer overflow, $1 billion prototype destroyed)
- Mars Climate Orbiter (1999)
  (destroyed in Mars orbit, mixup of units pound-force and newtons)
- Knight Capital Group Error in Ultra Short Time Trading (2012)
  (faulty deployment, repurposing of critical flag, $440 lost in 45 min on stock exchange)
- ...

Fun to read

`http://www.cs.tau.ac.il/~nachumd/verify/horror.html`
`https://en.wikipedia.org/wiki/List_of_software_bugs`

Proof

- proof can show absence of errors in design
- but proofs talk about a **design**, not a **real system**
- ⇒ testing and proving complement each other

“As far as the laws of mathematics refer to reality, they are not certain; and as far as they are certain, they do not refer to reality.”
— Albert Einstein

Mathematical vs. Formal Proof

Mathematical Proof

- informal, convince other mathematicians
- checked by community of domain experts
- subtle errors are hard to find
- often provide some new insight about our world
- often short, but require creativity and a brilliant idea

Formal Proof

- formal, rigorously use a logical formalism
- checkable by *stupid* machines
- very reliable
- often contain no new ideas and no amazing insights
- often long, very tedious, but largely trivial

We are interested in formal proofs in this lecture.

Detail Level of Formal Proof

In **Principia Mathematica** it takes 300 pages to prove 1+1=2.

This is nicely illustrated in **Logicomix - An Epic Search for Truth**.

Automated vs Manual (Formal) Proof

Fully Manual Proof

- very tedious one has to grind through many trivial but detailed proofs
- easy to make mistakes
- hard to keep track of all assumptions and preconditions
- hard to maintain, if something changes (see Ariane V)

Automated Proof

- amazing success in certain areas
- but still often infeasible for interesting problems
- hard to get insights in case a proof attempt fails
- even if it works, it is often not that automated
  - run automated tool for a few days
  - abort, change command line arguments to use different heuristics
  - run again and iterate till you find a set of heuristics that prove it fully automatically in a few seconds

Interactive Proofs

- combine strengths of manual and automated proofs
- many different options to combine automated and manual proofs
  - mainly check existing proofs (e. g. HOL Zero)
  - user mainly provides lemmata statements, computer searches proofs using previous lemmata and very few hints (e. g. ACL 2)
  - most systems are somewhere in the middle
- typically the human user
  - provides insights into the problem
  - structures the proof
  - provides main arguments
- typically the computer
  - checks proof
  - keeps track of all use assumptions
  - provides automation to grind through lengthy, but trivial proofs

Typical Interactive Proof Activities

- provide precise definitions of concepts
- state properties of these concepts
- prove these properties
  - human provides insight and structure
  - computer does book-keeping and automates simple proofs
- build and use libraries of formal definitions and proofs
  - formalisations of mathematical theories like
    - lists, sets, bags, ...
    - real numbers
    - probability theory
  - specifications of real-world artefacts like
    - processors
    - programming languages
    - network protocols
  - reasoning tools

There is a strong connection with programming. Lessons learned in Software Engineering apply.

Different Interactive Provers

- there are many different interactive provers, e. g.
  - Isabelle/HOL
  - Coq
  - PVS
  - HOL family of provers
  - ACL2
  - ...
- important differences
  - the formalism used
  - level of trustworthiness
  - level of automation
  - libraries
  - languages for writing proofs
  - user interface
  - ...

Which theorem prover is the best one? :-)

- there is no **best** theorem prover
- better question: Which is the **best one for a certain purpose**?
- important points to consider
  - existing libraries
  - used logic
  - level of automation
  - user interface
  - importance development speed versus trustworthiness
  - How familiar are you with the different provers?
  - Which prover do people in your vicinity use?
  - your personal preferences
  - ...

In this course we use the HOL theorem prover, because it is used by the TCS group.

Part II

Organisational Matters

Aims of this Course

Aims

- introduction to interactive theorem proving (ITP)
- being able to evaluate whether a problem can benefit from ITP
- hands-on experience with HOL
- learn how to build a formal model
- learn how to express and prove important properties of such a model
- learn about basic conformance testing
- use a theorem prover on a small project

Required Prerequisites

- some experience with functional programming
- knowing Standard ML syntax
- basic knowledge about logic (e. g. First Order Logic)

Dates

- Interactive Theorem Proving Course takes place in Period 4 of the academic year 2016/2017
- always in room 4523 or 4532
- each week

Mondays	10:15 - 11:45	lecture
Wednesdays	10:00 - 12:00	practical session
Fridays	13:00 - 15:00	practical session

- no lecture on Monday, 1st of May, instead on Wednesday, 3rd May
- last lecture: 12th of June
- last practical session: 21st of June
- 9 lectures, 17 practical sessions

Exercises

- after each lecture an exercise sheet is handed out
- work on these exercises alone, except if stated otherwise explicitly
- exercise sheet contains due date
  - usually 10 days time to work on it
  - hand in during practical sessions
  - lecture Monday $\longrightarrow$ hand in at latest in next week's Friday session
- main purpose: understanding ITP and learn how to use HOL
  - no detailed grading, just pass/fail
  - retries possible till pass
  - if stuck, ask me or one another
  - practical sessions intend to provide this opportunity

Practical Sessions

- very informal
- main purpose: work on exercises
  - I have a look and provide feedback
  - you can ask questions
  - I might sometimes explain things not covered in the lectures
  - I might provide some concrete tips and tricks
  - you can also discuss with each other
- attendance not required, but highly recommended
  - exception: session on 21st April
- only requirement: turn up long enough to hand in exercises
- **you need to bring your own computer**

Passing the ITP Course

- there is only a pass/fail mark
- to pass you need to
  - attend at least 7 of the 9 lectures
  - pass 8 of the 9 exercises

Communication

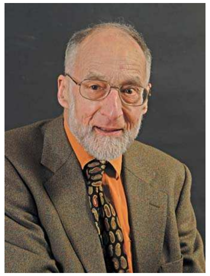
- we have the advantage of being a small group
- therefore we are flexible
- so please ask questions, even during lectures
- there are many shy people, therefore
  - anonymous checklist after each lecture
  - anonymous background questionnaire in first practical session
- further information is posted on **Interactive Theorem Proving Course** group on Group Web
- contact me (Thomas Tuerk) directly, e. g. via email `thomas@kth.se`

Part III

HOL 4 History and Architecture

LCF - Logic of Computable Functions

- **Standford LCF** 1971-72 by Milner et al.
- formalism devised by Dana Scott in 1969
- intended to reason about recursively defined functions
- intended for computer science applications
- strengths
  - powerful simplification mechanism
  - support for backward proof
- limitations
  - proofs need a lot of memory
  - fixed, hard-coded set of proof commands



Robin Milner
(1934 - 2010)

LCF - Logic of Computable Functions II

- Milner worked on improving LCF in Edinburgh
- research assistants
  - Lockwood Morris
  - Malcolm Newey
  - Chris Wadsworth
  - Mike Gordon
- **Edinburgh LCF** 1979
- introduction of **Meta Language** (ML)
- ML was invented to write proof procedures
- ML become an influential functional programming language
- using ML allowed implementing the **LCF approach**

LCF Approach

- implement an abstract datatype **thm** to represent theorems
- semantics of ML ensure that values of type thm can only be created using its interface
- interface is very small
  - predefined theorems are axioms
  - function with result type theorem are inferences
- $\Longrightarrow$ However you create a theorem, it is valid.
- together with similar abstract datatypes for types and terms, this forms the **kernel**

LCF Approach II

Modus Ponens Example

Inference Rule

$$\frac{\Gamma \vdash a \Rightarrow b \qquad \Delta \vdash a}{\Gamma \cup \Delta \vdash b}$$

SML function

```
val MP : thm -> thm -> thm
```

$\mathtt{MP}(\Gamma \vdash a \Rightarrow b)(\Delta \vdash a) = (\Gamma \cup \Delta \vdash b)$

- very trustworthy — only the small kernel needs to be trusted
- efficient — no need to store proofs

Easy to extend and automate

However complicated and potentially buggy your code is, if a value of type theorem is produced, it has been created through the small trusted interface. Therefore the statement really holds.

LCF Style Systems

There are now many interactive theorem provers out there that use an approach similar to that of Edinburgh LCF.

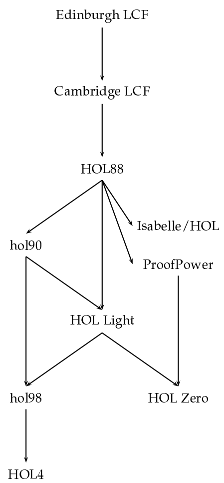
- HOL family
  - HOL theorem prover
  - HOL Light
  - HOL Zero
  - Proof Power
  - ...
- Isabelle
- Nuprl
- Coq
- ...

History of HOL

- 1979 Edinburgh LCF by Milner, Gordon, et al.
- 1981 Mike Gordon becomes lecturer in Cambridge
- 1985 Cambridge LCF
  - Larry Paulson and Gèrard Huet
  - implementation of ML compiler
  - powerful simplifier
  - various improvements and extensions
- 1988 HOL
  - Mike Gordon and Keith Hanna
  - adaption of Cambridge LCF to classical higher order logic
  - intention: hardware verification
- 1990 HOL90
  reimplementation in SML by Konrad Slind at University of Calgary
- 1998 HOL98
  implementation in Moscow ML and new library and theory mechanism
- since then HOL Kananaskis releases, called informally **HOL 4**

Family of HOL

- **ProofPower**
  commercial version of HOL88 by Roger Jones, Rob Arthan et al.
- **HOL Light**
  lean CAML / OCaml port by John Harrison
- **HOL Zero**
  trustworthy proof checker by Mark Adams
- **Isabelle**
  - 1990 by Larry Paulson
  - meta-theorem prover that supports multiple logics
  - however, mainly HOL used, ZF a little
  - nowadays probably the most widely used HOL system
  - originally designed for software verification



Part IV

HOL's Logic

HOL Logic

- the HOL theorem prover uses a version of classical **h**igher **o**rder **l**ogic: classical higher order predicate calculus with terms from the typed lambda calculus (i. e. simple type theory)
- this sounds complicated, but is intuitive for SML programmers
- (S)ML and HOL logic designed to fit each other
- if you understand SML, you understand HOL logic

HOL = functional programming + logic

Ambiguity Warning

The acronym *HOL* refers to both the *HOL interactive theorem prover* and the *HOL logic* used by it. It's also a common abbreviation for *higher order logic* in general.

Types

- SML datatype for types
  - **Type Variables** (`'a`, α, `'b`, β, ...)
    Type variables are implicitly universally quantified. Theorems containing type variables hold for all instantiations of these. Proofs using type variables can be seen as proof schemata.
  - **Atomic Types** (`c`)
    Atomic types denote fixed types. Examples: `num`, `bool`, `unit`
  - **Compound Types** ($(\sigma_1, \ldots, \sigma_n) op$)
    op is a **type operator** of arity *n* and $\sigma_1, \ldots, \sigma_n$ **argument types**. Type operators denote operations for constructing types.
    Examples: `num list` or `'a # 'b`.
  - **Function Types** ($\sigma_1 \to \sigma_2$)
    $\sigma_1 \to \sigma_2$ is the type of **total** functions from σ_1 to σ_2.
- types are never empty in HOL, i. e. for each type at least one value exists
- all HOL functions are total

Terms

- SML datatype for terms
  - **Variables** (`x`, `y`, ...)
  - **Constants** (`c`, ...)
  - **Function Application** (`f a`)
  - **Lambda Abstraction** (`\x. f x` or $\lambda x.\ fx$)
    Lambda abstraction represents anonymous function definition.
    The corresponding SML syntax is `fn x => f x`.
- terms have to be well-typed
- same typing rules and same type-inference as in SML take place
- terms very similar to SML expressions
- notice: predicates are functions with return type `bool`, i. e. no distinction between functions and predicates, terms and formulae

Terms II

HOL term	**SML expression**	**type HOL / SML**
0	0	`num` / `int`
`x:'a`	`x:'a`	variable of type `'a`
`x:bool`	`x:bool`	variable of type `bool`
`x + 5`	`x + 5`	applying function `+` to `x` and `5`
`\x. x + 5`	`fn x => x + 5`	anonymous (a. k. a. inline) function of type `num -> num`
`(5, T)`	`(5, true)`	`num # bool` / `int * bool`
`[5;3;2]++[6]`	`[5,3,2]@[6]`	`num list` / `int list`

Free and Bound Variables / Alpha Equivalence

- the lambda-expression $\lambda x.\ t$ is said to **bind** the variables x in term t
- variables that are guarded by a lambda expression are called **bound**
- all other variables are **free**
- Example: x is free and y is bound in $(x = 5) \wedge (\lambda y.\ (y < x))\ 3$
- the names of bound variables are unimportant semantically
- two terms are called **alpha-equivalent** iff they differ only in the names of bound variables
- Example: $\lambda x.\ x$ and $\lambda y.\ y$ are alpha-equivalent
- Example: x and y are not alpha-equivalent

Theorems

- theorems are of the form $\Gamma \vdash p$ where
  - Γ is a set of hypothesis
  - p is the conclusion of the theorem
  - all elements of Γ and p are formulae, i.e. terms of type `bool`
- $\Gamma \vdash p$ records that using Γ the statement p **has been** proved
- notice difference to logic: there it means **can be** proved
- the proof itself is not recorded
- theorems can only be created through a small interface in the **kernel**

HOL Light Kernel

- the HOL kernel is hard to explain
  - for historic reasons some concepts are represented rather complicated
  - for speed reasons some derivable concepts have been added
- instead consider the HOL Light kernel, which is a cleaned-up version
- there are two predefined constants
  - `= : 'a -> 'a -> bool`
  - `@ : ('a -> bool) -> 'a`
- there are two predefined types
  - `bool`
  - `ind`
- the meaning of these types and constants is given by inference rules and axioms

HOL Light Inferences I

$$\frac{}{\vdash t = t}\ \mathrm{REFL}$$

$$\frac{\Gamma \vdash s = t \qquad \Delta \vdash t = u}{\Gamma \cup \Delta \vdash s = u}\ \mathrm{TRANS}$$

$$\frac{\Gamma \vdash s = t \qquad \Delta \vdash u = v \qquad \textit{types fit}}{\Gamma \cup \Delta \vdash s(u) = t(v)}\ \mathrm{COMB}$$

$$\frac{\Gamma \vdash s = t \qquad x \textit{ not free in } \Gamma}{\Gamma \vdash \lambda x.\ s = \lambda x.\ t}\ \mathrm{ABS}$$

$$\frac{}{\vdash (\lambda x.\ t)x = t}\ \mathrm{BETA}$$

$$\frac{}{\{p\} \vdash p}\ \mathrm{ASSUME}$$

HOL Light Inferences II

$$\frac{\Gamma \vdash p \Leftrightarrow q \qquad \Delta \vdash p}{\Gamma \cup \Delta \vdash q}\ \mathrm{EQ\_MP}$$

$$\frac{\Gamma \vdash p \qquad \Delta \vdash q}{(\Gamma - \{q\}) \cup (\Delta - \{p\}) \vdash p \Leftrightarrow q}\ \mathrm{DEDUCT\_ANTISYM\_RULE}$$

$$\frac{\Gamma[x_1, \ldots, x_n] \vdash p[x_1, \ldots, x_n]}{\Gamma[t_1, \ldots, t_n] \vdash p[t_1, \ldots, t_n]}\ \mathrm{INST}$$

$$\frac{\Gamma[\alpha_1, \ldots, \alpha_n] \vdash p[\alpha_1, \ldots, \alpha_n]}{\Gamma[\gamma_1, \ldots, \gamma_n] \vdash p[\gamma_1, \ldots, \gamma_n]}\ \mathrm{INST\_TYPE}$$

HOL Light Axioms and Definition Principles

- 3 axioms needed

ETA_AX	$(\lambda x.\ t\ x) = t$
SELECT_AX	$P\ x \Longrightarrow P((@)P))$
INFINITY_AX	predefined type `ind` is infinite

- definition principle for constants
  - constants can be introduced as abbreviations
  - constraint: no free vars and no new type vars
- definition principle for types
  - new types can be defined as non-empty subtypes of existing types
- both principles
  - lead to conservative extensions
  - preserve consistency

HOL Light derived concepts

Everything else is derived from this small kernel.

$$\begin{array}{lll}
T & =_{def} & (\lambda p.\ p) = (\lambda p.\ p) \\
\wedge & =_{def} & \lambda p\, q.\ (\lambda f.\ f\ p\ q) = (\lambda f.\ f\ T\ T) \\
\Longrightarrow & =_{def} & \lambda p\, q.\ (p \wedge q \Leftrightarrow p) \\
\forall & =_{def} & \lambda P.\ (P = \lambda x.\ T) \\
\exists & =_{def} & \lambda P.\ (\forall q.\ (\forall x.\ P(x) \Longrightarrow q) \Longrightarrow q) \\
\ldots & &
\end{array}$$

Multiple Kernels

- Kernel defines abstract datatypes
- one does not need to look at the internal implementation
- therefore, easy to exchange
- there are at least 3 different kernels for HOL
  - standard kernel (de Bruijn indices)
  - experimental kernel (name / type pairs)
  - OpenTheory kernel (for proof recording)

HOL Logic Summary

- HOL theorem prover uses classical higher order logic
- HOL logic is very similar to SML
  - syntax
  - type system
  - type inference
- HOL theorem prover very trustworthy because of LCF approach
  - there is a small kernel
  - proofs are not stored explicitly
- you don't need to know the details of the kernel
- usually one works at a much higher level of abstraction