

► Goal is to show formal semantics is not just documentation

- ► Goal is to show formal semantics is not just documentation
 - can run the Language Reference Manual (LRM)

- ► Goal is to show formal semantics is not just documentation
 - can run the Language Reference Manual (LRM)
- Correctness primary, efficiency secondary

- Goal is to show formal semantics is not just documentation
 - can run the Language Reference Manual (LRM)
- Correctness primary, efficiency secondary
 - but need sufficient efficiency!

- Goal is to show formal semantics is not just documentation
 - can run the Language Reference Manual (LRM)
- Correctness primary, efficiency secondary
 - but need sufficient efficiency!
- Programming methodology, not new verification algorithms

- Goal is to show formal semantics is not just documentation
 - can run the Language Reference Manual (LRM)
- Correctness primary, efficiency secondary
 - but need sufficient efficiency!
- Programming methodology, not new verification algorithms
 - EDA tools with theorem prover inside (c.f. PROSPER)

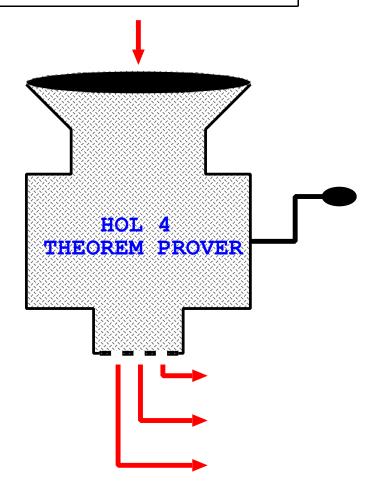
Accellera's PSL (formerly IBM's Sugar 2.0)

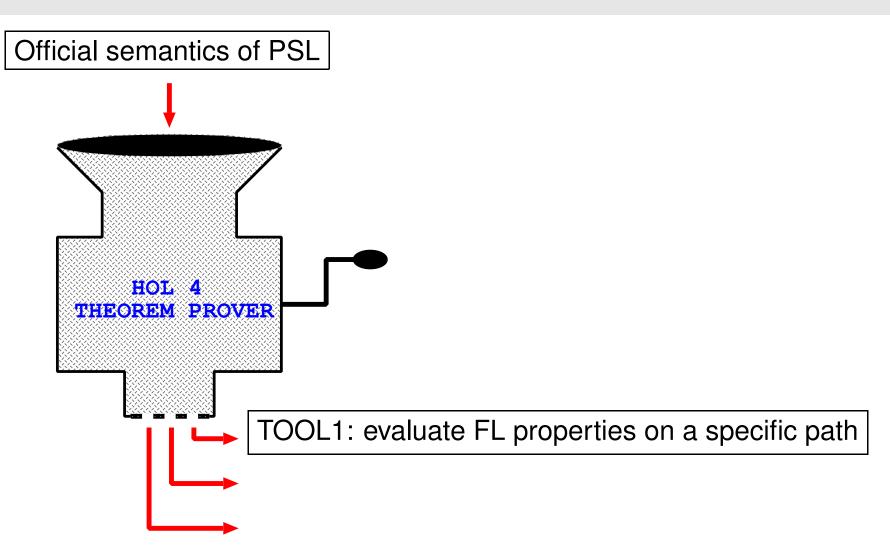
- PSL is a property specification language combining
 - boolean expressions

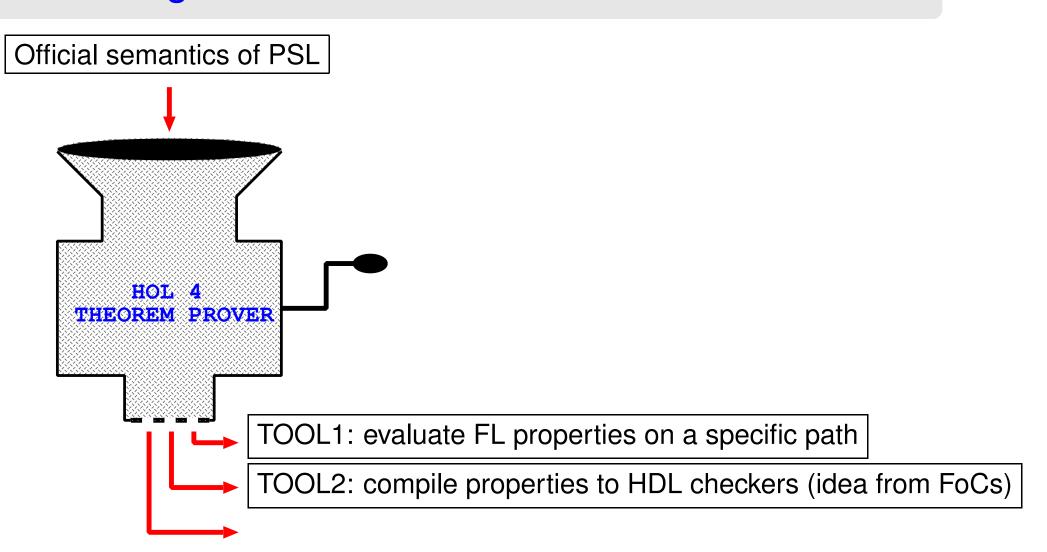
(Verilog syntax)

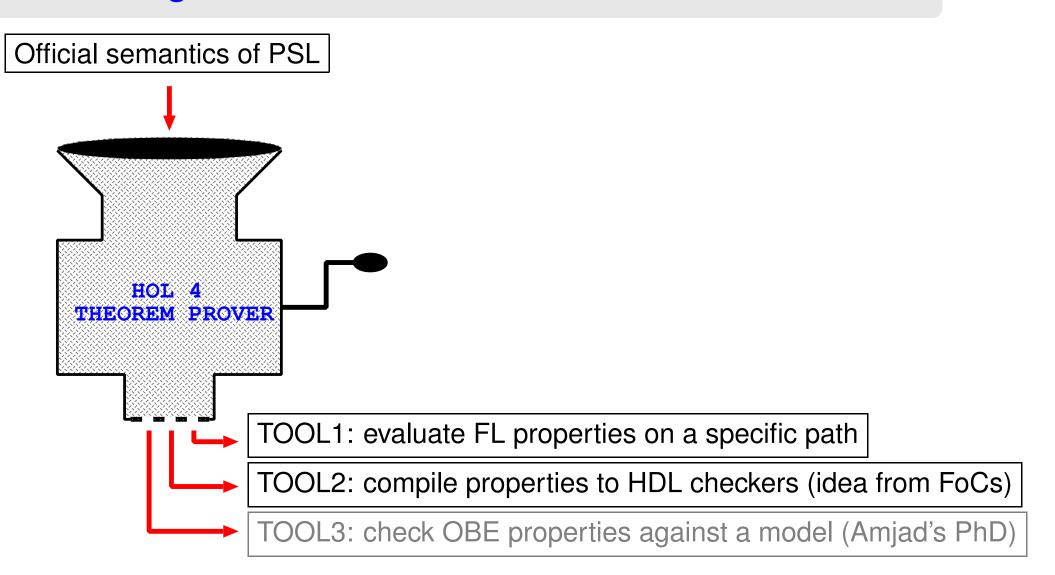
- patterns (Sequential Extended Regular Expressions SEREs)
- LTL formulas (Foundation language FL)
- CTL formulas (Optional Branching Extension OBE)
- Designed both for model checking and simulation testbenches
- Intended to be the industry standard

Official semantics of PSL









Tools use standard algorithms

- ▶ TOOL1: semantic calculator
 - match regexps using automata; evaluate formulas recursively
 - automata constructed and executed by proof inside HOL
- ▶ TOOL2: checker compiler
 - compile regexps to automata, then 'pretty print' to HDL (Verilog)
 - treatment of formulas incomplete and ad hoc
- ➤ TOOL3: symbolic model checker
 - ullet classical McMillan-style μ -calculus checker
 - uses BDD representation judgements to link HOL terms to BDDs
 - see Gordon (TPHOLs2001), Amjad (TPHOLs2003)
- No new algorithms, but maybe a new kind of logic programming

Our theorem proving infrastructure (HOL)



- \triangleright Standard ML infrastructure to interactively prove $\vdash t$
 - t is a term of higher order logic
 - proof is 'fully-expansive' a sequence of primitive inference steps
- Logic is typed
 - type system supports user-defined datatypes
 - example: define types of PSL expressions, SEREs and formulas
- Contains the usual proof tools
 - simplifier (rewriter)
 - decision procedures for subsets of natural numbers, integers, reals
 - first order reasoners (inspired by Isabelle)

Heroic proofs versus logic programming



- ► Theorem proving often associated with heroic proofs
 - e.g. Gödel's theorem (Shankar), relative consistency of AC (Paulson)
- ▶ We are not doing heroic proofs, but a kind of logic programming
 - computation by deduction
- ► HOL has a relatively fast call-by-value symbolic evaluator EVAL
 - by Bruno Barras using Coq technology (explicit substitutions)
 - doesn't compete with ACL2 or PVS ground evaluators (or C, C++)
 - runs ARM6 microarchitecture at a few seconds per instruction
 - key tool for our PSL evalutor

Semantics of boolean expressions (PSL in red, HOL in blue)

$$(s \models p = p \in s) \land (s \models \neg b = \neg (s \models b)) \land (s \models b_1 \land b_2 = s \models b_1 \land s \models b_2)$$

Fragment of semantics of formulas

```
(w \models b = |w| > 0 \land w_0 \models b) \land
(w \models f_1 \land f_2 = w \models f_1 \land w \models f_2) \land
(w \models X! f = |w| > 1 \land w^1 \models f)
```

Examples of rewriting and evaluation:

 \vdash

 \vdash

 \vdash

 \vdash

Semantics of boolean expressions (PSL in red, HOL in blue)

$$(s \models p = p \in s) \land (s \models \neg b = \neg (s \models b)) \land (s \models b_1 \land b_2 = s \models b_1 \land s \models b_2)$$

Fragment of semantics of formulas

```
(w \models b = |w| > 0 \land w_0 \models b) \land 
(w \models f_1 \land f_2 = w \models f_1 \land w \models f_2) \land 
(w \models X! f = |w| > 1 \land w^1 \models f)
```

```
 \vdash w \models p \land X! \ f = (|w| > 0 \land w_0 \models p) \land |w| > 1 \land w^1 \models f   \vdash
```

Semantics of boolean expressions (PSL in red, HOL in blue)

$$(s \models p = p \in s) \land (s \models \neg b = \neg (s \models b)) \land (s \models b_1 \land b_2 = s \models b_1 \land s \models b_2)$$

Fragment of semantics of formulas

```
(w \models b = |w| > 0 \land w_0 \models b) \land
(w \models f_1 \land f_2 = w \models f_1 \land w \models f_2) \land
(w \models X! f = |w| > 1 \land w^1 \models f)
```

```
\vdash w \models p \land X! f = (|w| > 0 \land w_0 \models p) \land |w| > 1 \land w^1 \models f
\vdash [s_0]w \models p \land X! f = s_0 \models p \land |w| + 1 > 1 \land w \models f
\vdash
```

Semantics of boolean expressions (PSL in red, HOL in blue)

$$(s \models p = p \in s) \land (s \models \neg b = \neg (s \models b)) \land (s \models b_1 \land b_2 = s \models b_1 \land s \models b_2)$$

Fragment of semantics of formulas

$$(w \models b = |w| > 0 \land w_0 \models b) \land$$

$$(w \models f_1 \land f_2 = w \models f_1 \land w \models f_2) \land$$

$$(w \models X! f = |w| > 1 \land w^1 \models f)$$

```
\vdash w \models p \land X! f = (|w| > 0 \land w_0 \models p) \land |w| > 1 \land w^1 \models f
\vdash [s_0]w \models p \land X! f = s_0 \models p \land |w| + 1 > 1 \land w \models f
\vdash s_0 s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8 s_9 \models p \land X! f = s_0 \models p \land s_1 s_2 s_3 s_4 s_5 s_6 s_7 s_8 s_9 \models f
\vdash
```

Semantics of boolean expressions (PSL in red, HOL in blue)

$$(s \models p = p \in s) \land (s \models \neg b = \neg (s \models b)) \land (s \models b_1 \land b_2 = s \models b_1 \land s \models b_2)$$

Fragment of semantics of formulas

$$(w \models b = |w| > 0 \land w_0 \models b) \land$$

$$(w \models f_1 \land f_2 = w \models f_1 \land w \models f_2) \land$$

$$(w \models X! f = |w| > 1 \land w^1 \models f)$$

```
 \vdash w \models p \land X! \ f = (|w| > 0 \land w_0 \models p) \land |w| > 1 \land w^1 \models f 
 \vdash [s_0]w \models p \land X! \ f = s_0 \models p \land |w| + 1 > 1 \land w \models f 
 \vdash s_0s_1s_2s_3s_4s_5s_6s_7s_8s_9 \models p \land X! \ f = s_0 \models p \land s_1s_2s_3s_4s_5s_6s_7s_8s_9 \models f 
 \vdash \{a\}\{a,b\}\{b\} \models a \land X! \ b = \mathsf{T}
```

Parts of semantics require reformulation for execution

► LRM semantics of the until-operator not directly executable

$$w \models [f_1 \ U \ f_2] = \exists k \in [0 ... |w|). \ w^k \models f_2 \land \forall j \in [0 ... k). \ w^j \models f_1$$

Standard reformulation makes it directly executable

$$\vdash w \models [f_1 \ U \ f_2] = |w| > 0 \land (w \models f_2 \lor w \models f_1 \land w^1 \models [f_1 \ U \ f_2])$$

▶ If f_1 , f_2 are boolean expressions and the path is arbitrary of length 5:

```
\vdash s_0 s_1 s_2 s_3 s_4 \models [b_1 \ U \ b_2] =
s_0 \models b_2 \lor
s_0 \models b_1 \land (s_1 \models b_2 \lor s_1 \models b_1 \land
(s_2 \models b_2 \lor s_2 \models b_1 \land (s_3 \models b_2 \lor s_3 \models b_1 \land s_4 \models b_2)))
```

Matching regular expressions

Semantics of PSL SEREs is self-explanatory

```
 (w \models b \qquad \qquad = (|w| = 1) \land w_0 \models b) \qquad \qquad \land 
 (w \models r_1; \ r_2 \qquad \qquad = \exists \ w_1 w_2. \ (w = w_1 w_2) \land w_1 \models r_1 \land w_2 \models r_2) \qquad \land 
 (w \models r_1 : r_2 \qquad \qquad = \exists \ w_1 w_2 l. \ (w = w_1 [l] w_2) \land w_1 [l] \models r_1 \land [l] w_2 \models r_2) \land 
 (w \models \{r_1\} \mid \{r_2\} \qquad = \ w \models r_1 \lor w \models r_2) \qquad \qquad \land 
 (w \models \{r_1\} \& \& \{r_2\} = \ w \models r_1 \land w \models r_2) \qquad \qquad \land 
 (w \models r[*] \qquad \qquad = \exists \ w list. \ (w = \mathsf{Concat} \ w list) \land \mathsf{Every}(\lambda \ w. \ w \models r) w list)
```

Make executable by proving

```
\vdash \forall w \ r. \ w \models r = \operatorname{amatch}(\operatorname{sere2regexp}(r))w
```

where:

- sere2regexp converts a SERE to a HOL regular expression
- amatch is an executable matcher for regular expressions

Suffix implication $\{r\}(f)$

Semantics is:

$$w \models \{r\}(f) = \forall j \in [0... |w|). w^{0,j} \models r \Rightarrow w^j \models f$$

► Have defined an efficient executable function acheck so that, for example:

```
acheck r f [x_0; x_1; x_2; x_3] =
(\mathsf{amatch} \ r \ [x_0] \Rightarrow f[x_0; x_1; x_2; x_3]) \land \\ (\mathsf{amatch} \ r \ [x_0; x_1] \Rightarrow f[x_1; x_2; x_3]) \land \\ (\mathsf{amatch} \ r \ [x_0; x_1; x_2] \Rightarrow f[x_2; x_3]) \land \\ (\mathsf{amatch} \ r \ [x_0; x_1; x_2; x_3] \Rightarrow f[x_3])
```

Then proved

```
\vdash \forall w \ r \ f. \ w \models \{r\}(f) = \text{acheck}(\text{sere2regexp}(r))(\lambda x. \ x \models f)w
```

Rewrite with this, then execute

Strong suffix implication $\{r_1\} \mapsto \{r_2\}!$

> Semantics is:

$$w \models \{r_1\} \mapsto \{r_2\}! = \forall j \in [0 ... |w|).w^{0,j} \models r_1 \Rightarrow \exists k \in [j ... |w|).w^{j,k} \models r_2$$

Reduced to suffix implication by proving

$$\vdash \ \forall w \ r_1 \ r_2. \ w \models \{r_1\} \mapsto \{r_2\}! = w \models \{r_1\}(\neg \{r_2\}(\mathsf{F}))$$

Rewrite with this, then execute

Weak suffix implication $\{r_1\} \mapsto \{r_2\}$

> Semantics is:

```
w \models \{r_1\} \mapsto \{r_2\} = 
\forall j \in [0 ... |w|).
w^{0,j} \models r_1 \Rightarrow (\exists k \in [j ... |w|).w^{j,k} \models r) \lor (\forall k \in [j ... |w|).\exists w'.w^{j,k}w' \models r_2)
```

▶ Have added a special regular expression Prefix(r) to HOL (not to PSL)

```
\vdash \forall r \ w. \ w \models \mathsf{Prefix}(r) = \exists w'. \ w \ w' \models r
```

- ightharpoonup Execution of $w \models \mathsf{Prefix}(r)$ uses Dijkstra's algorithm
- Have proved:

```
 \begin{array}{c} \vdash \ \forall w \ r_1 \ r_2. \\ w \models \{r_1\} \mapsto \{r_2\} = \\ \text{acheck(sere2regexp} \ r_1) \\ (\lambda \, x. \ x \models \neg \{r_2\}(\mathsf{F}) \ \lor \ \text{amatch (Prefix (sere2regexp} \ r_2))} \ x) \ w \end{array}
```

Rewrite with this, then execute

Remaining formulas: aborts and clocking

Semantics of abort formulas:

```
w \models f \ abort \ b = w \models f \ \lor \ w \models b \lor \exists j \in [1 ... |w|). \ \exists w'. \ w^j \models b \land w^{0,j-1}w' \models f
```

- $\exists w'$ needs a reachability algorithm
- have implemented a partial method
- awaiting new abort semantics before attempting complete solution
- ▶ Clocked formulas f@c, f@c! can be translated to unclocked formulas
 - translation to unclocked formulas is by a recursive function
 - can be directly executed

Clocking

- ▶ LRM defines $w \stackrel{\mathcal{C}}{\models} r$ and $w \stackrel{\mathcal{C}}{\models} f$ for arbitrary clock c
 - clocks c are arbitrary boolean expressions
 - top level default clock is T
- Semantics of clocked SEREs

$$w \stackrel{c}{\models} r@c_1 = \exists i \in [0 ... |w|). \ w^{0,i} \stackrel{\mathsf{T}}{\models} \neg c_1[*]; \ c_1 \wedge w^i \stackrel{c_1}{\models} r$$

Semantics of clocked formulas

$$w \stackrel{c}{\models} f @c_1! = \exists i \in [0 ... |w|). \ w^{0,i} \stackrel{\mathsf{T}}{\models} \neg c_1[*]; \ c_1 \wedge w^i \stackrel{c}{\models} f$$

 \triangleright Execute by rewriting with function \mathcal{T}^{T} and then the theorems:

$$\vdash \ \forall r \ w. \ w \models^{\mathsf{T}} r = w \models \mathcal{T}^{\mathsf{T}}(r)$$

$$\vdash \forall f \ w. \ w \models^{\mathsf{T}} f = w \models \mathcal{T}^{\mathsf{T}}(f)$$

Example

▶ PSL Reference Manual Example 2, page 45

time	0	1	2	3	4	5	6	7	8	9
clk1 a b c clk2	0 0 0 0 1 1	1 0 0 0	0 0 0	1 1 0 0	0	1 1	0 0 0 0 1 1	1 0 1 0 0	0 0 1 0	1 0 0 0

▶ Define w to be this path, so w is :

```
{c,clk2}{clk1}{}{clk1,a,clk2}{a}{clk1,a,b,c}{c,clk2}{clk1,b}{b}{clk1,clk2}
```

- **Example** uses weak clocking defined by: $f@c = \neg(\neg f@c!)$
- Evaluation yields
 - $\vdash \mathbf{w}^{6} \models^{\mathsf{T}} (c \land X! [a \ U \ b] @(clk_{1} \lor clk_{2})) @(clk_{1} \lor clk_{2}) = \mathsf{T}$
 - $\vdash \mathbf{w}^i \models^\mathsf{T} (c \land X! [a \ U \ b] @(clk_1 \lor clk_2)) @(clk_1 \lor clk_2) = \mathsf{F} \text{ (if } i \neq 6)$

SML convenient for scripting combinations of evaluations

Example: use SML map function to generate

Easy to evaluate SEREs and formulas on all subpaths of a path

Uses of TOOL1 (calculating $w \models f$ from semantics)

- Teaching and learning tool for exploring semantics
- Checking one has the right property before using it in verification
 - Post simulation analysis (path is generated by simulator)
 - compare with "TransEDA VN-Property" property checker and analyzer
 - our tools much slower but not necessary too slow!
 - guaranteed PSL compliant by construction: golden reference

TOOL2: Compile the semantics to checkers

- Idea pinched from IBM FoCs project
- ▶ A defined operator: $\forall r. never(r) = \{T[*]; r\} \mapsto \{F\}$
- ightharpoonup Example property: $never(\neg stob_reQ \land btos_ack; stob_reQ)$
- Use semantics to generate a Verilog checker

```
module Checker (StoB_REQ, BtoS_ACK, BtoR_REQ, RtoB_ACK);
input StoB_REQ, BtoS_ACK, BtoR_REQ, RtoB_ACK;
reg [1:0] state;
initial state = 0;
always @ (StoB_REQ or BtoS_ACK or BtoR_REQ or RtoB_ACK)
begin
 $display ("Checker: state = %0d", state);
 case (state)
  0: if (StoB_REQ) state = 1; else if (BtoS_ACK) state = 2; else state = 1;
  1: if (StoB_REQ) state = 1; else if (BtoS_ACK) state = 2; else state = 1;
  2: if (StoB_REQ) state = 3; else if (BtoS_ACK) state = 2; else state = 1;
  3: begin $display ("Checker: property violated!"); $finish; end
  default: begin $display ("Checker: unknown state"); $finish; end
 endcase
end
endmodule
```

Example of how the checker works and is justified

The following theorem is first proved

```
|w| = \infty \implies w \models never(r) = \forall n. \neg amatch (sere2regexp T[*]; r)(w^{0,n})
```

- ▶ Thus there's an error if amatch (sere2regexp T[*]; $r)(w^{0,n})$ is ever true
- ▶ Generate a DFA from sere2regexp T[*]; r
- So far everything is by proof, so correct by construction
- Final step is to pretty print checker into HDL (Verilog)
 - this may introduce errors
 - no formal semantics of Verilog : (
- Only have 'proof of concept' for checkers: more work to cover all formulas

Conclusions

- ► Two tools: semantic calculator and checker generator
- Correct by construction
- More work needed (especially for checkers)
 - Illustrates new kind of logic programming using a theorem prover
 - prototyping standards compliant tools
 - theorem proving is slow
 - maybe OK for some industrial strength performance-non-critical tools
- Possible application: generate OVL checkers from PSL specifications

Conclusions



- ► Two tools: semantic calculator and checker generator
- Correct by construction
- More work needed (especially for checkers)
 - Illustrates new kind of logic programming using a theorem prover
 - prototyping standards compliant tools
 - theorem proving is slow but not necessarily too slow
 - maybe OK for some industrial strength performance-non-critical tools
- Possible application: generate OVL checkers from PSL specifications

Conclusions



- ► Two tools: semantic calculator and checker generator
- Correct by construction
- More work needed (especially for checkers)
- Illustrates new kind of logic programming using a theorem prover
 - prototyping standards compliant tools
 - theorem proving is slow but not necessarily too slow
 - maybe OK for some industrial strength performance-non-critical tools
- > Possible application: generate OVL checkers from PSL specifications
- THE END

ADDITIONAL SLIDES ON HOL



The HOL system

- Versions of the HOL system:

 - HOL88 from Cambridge
 HOL90 from Calgary and Bell Labs
 HOL98 from Cambridge, Glasgow and Utah.
 HOL 4 open source project at SourceForge

The HOL system

- Versions of the HOL system:

 - HOL88 from Cambridge
 HOL90 from Calgary and Bell Labs
 HOL98 from Cambridge, Glasgow and Utah.
 HOL 4 open source project at SourceForge

	Current team	h	nol.sf.	net
--	--------------	---	---------	-----

Developer	Role/Position	Location
Anthony Fox Peter Homeier Hasan Amjad	Developer Developer Developer	UK USA UK
Joe Hurd Ken Friis Larsen Keith Wansbrough Michael Norrish Mike Gordon Konrad Slind	Developer Advisor/Mentor/Consultant Developer Project Manager Developer Project Manager Project Manager	UK Denmark UK Australia UK USA

The HOL system

- Versions of the HOL system:

 - HOL88 from Cambridge
 HOL90 from Calgary and Bell Labs
 HOL98 from Cambridge, Glasgow and Utah.
 HOL 4 open source project at SourceForge

Current team		hol.sf.net
	Current team	Current team

Developer	Role/Position	Location
Anthony Fox	Developer	UK
Peter Homeier	Developer	USA
Hasan Amjad	Developer	UK
Joe Hurd	Developer	UK
Ken Friis Larsen	Advisor/Mentor/Consultant	Denmark
Keith Wansbrough	Developer	UK
Michael Norrish	Project Manager	Australia
Mike Gordon	Developer	UK
Konrad Slind	Project Manager	USA

No longer managed from Cambridge

New tools (some here, some coming soon)



- New theorem proving tactics
 - ordered resolution and paramodulation for equality reasoning
 - time-sliced combinations of resolution and model elimination
- ▶ New decision procedure for full Presburger arithmetic
 - Pugh's "Omega Test"
- Improved support for emulating predicate subtypes
 - PVS is still better :-(
- Fully-expansive model checking
 - CTL checking as proof in representation judgement calculus
- Tools for 'boolification' to encode for BDD and SAT
 - automatically generate encoders/decoders from datatype definition
 - automatically generate bitvector versions of function definitions

Some recent or current projects



- Verification of AES (Rijndael) and others (Serpent, MARS, Twofish, RC6)
 - synergy between symbolic execution and proof
 - Slind and students (Utah)
- Memory models
 - general model applied to Java threads
 - Slind/Gopalakrishnan and students (Utah)
- ARM processor verification
 - programmers view of ARM6 equivalent to pipelined microarchitecture
 - Fox (Cambridge), Birtwistle and students (Leeds) and ARM
 - future work is ESL verification using ARM model
- Verification of probabilistic algorithms
 - Miller-Rabin probabilistic primality test
 - Hurd (Cambridge)
- Mechanised semantics of realistic networking (UDP)
 - validate operational semantics of network programming protocols
 - Sewell/Wansbrough & Norrish (Cambridge & Australia)