

Half talk: Translation Validation for seL4

Thomas Sewell Magnus Myreen

NICTA & Cambridge

10 May 2013



Australian Government
Department of Broadband,
Communications and the Digital Economy
Australian Research Council

NICTA Funding and Supporting Members and Partners



Australian
National
University



UNSW
THE UNIVERSITY OF NEW SOUTH WALES



Trade &
Investment



QUT



Victoria
The Place to Be



Queensland
Government



THE UNIVERSITY OF
SYDNEY



Queensland
Government

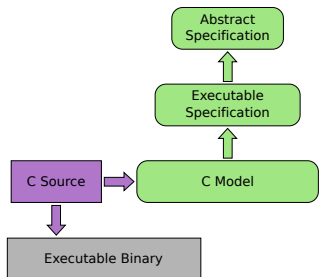


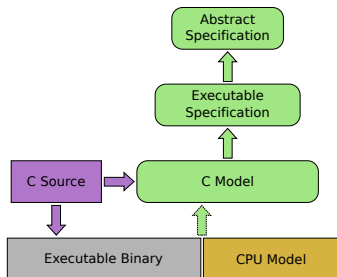
Griffith
UNIVERSITY



THE UNIVERSITY OF
QUEENSLAND
AUSTRALIA

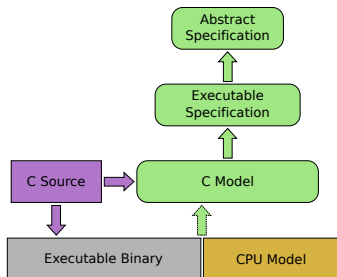
We want to link the seL4 proofs down to the binary, giving the proof deeper foundations.





We want to link the seL4 proofs down to the binary, giving the proof deeper foundations.

We link to the Cambridge ARM model, which is extensively validated. This brings our theory about as close to the real world as we can go.



We want to link the seL4 proofs down to the binary, giving the proof deeper foundations.

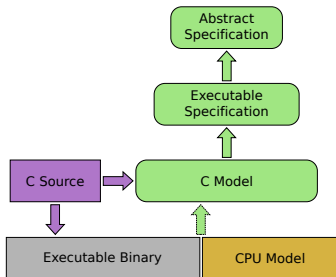
We link to the Cambridge ARM model, which is extensively validated. This brings our theory about as close to the real world as we can go.

This guards against the compiler being broken, the C semantics being wrong, or the standard being weak.

Comparisons

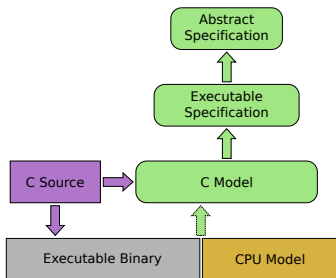


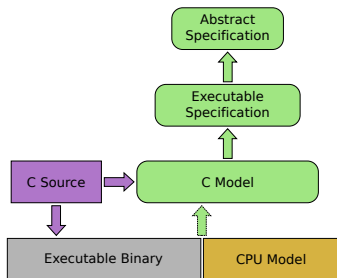
There are similarities to previous really big projects like Verisoft.



There are similarities to previous really big projects like Verisoft.

There is also a clear difference: this proof is happening incrementally.



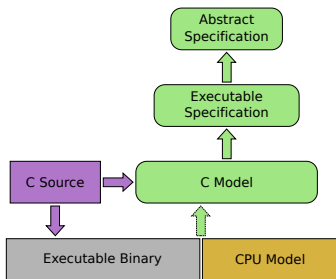


There are similarities to previous really big projects like Verisoft.

There is also a clear difference: this proof is happening incrementally.

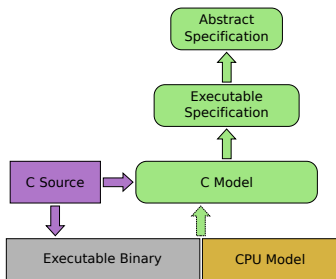
Also note this is only half the binary verification issue for seL4.

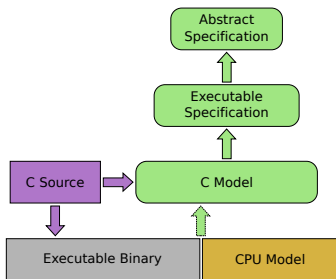
This is **translation validation**, a form of **refinement**.



This is **translation validation**, a form of **refinement**.

Similar to certified compilation, certifying compilation, binary verification and proof carrying code.





This is **translation validation**, a form of **refinement**.

Similar to certified compilation, certifying compilation, binary verification and proof carrying code.

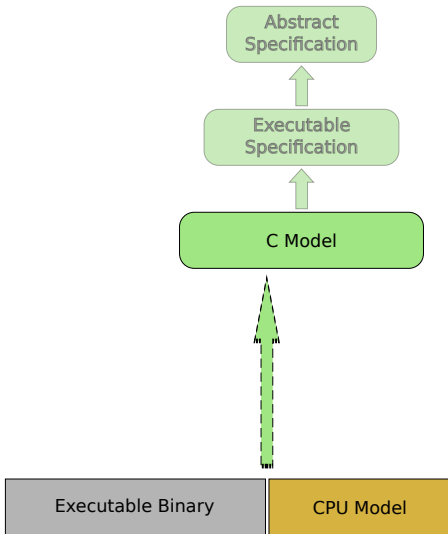
There's a lot of other work in this space. All that really distinguishes us is our motivation.

Motivation: We care about getting a result for one system and proof.

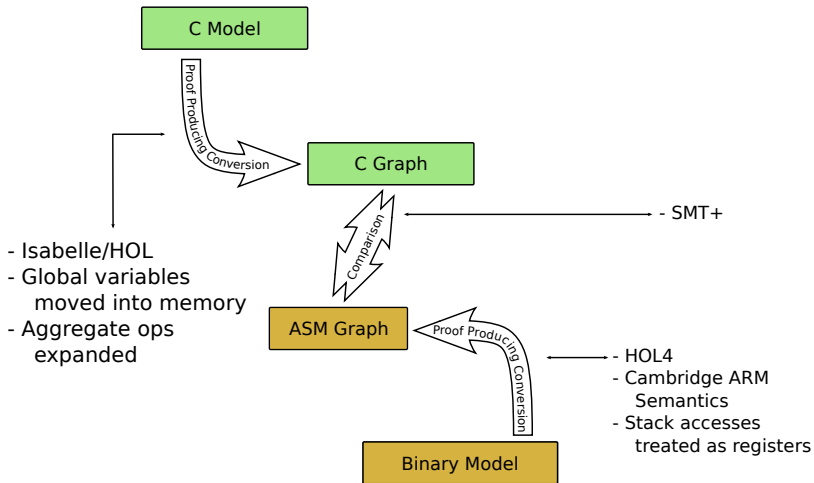
Motivation: We care about getting a result for one system and proof.
Period.

We don't care about performance, coverage of the C language or of C compiler optimisations. We don't care about gcc.

Approach



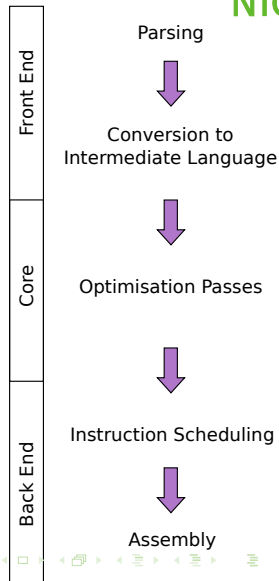
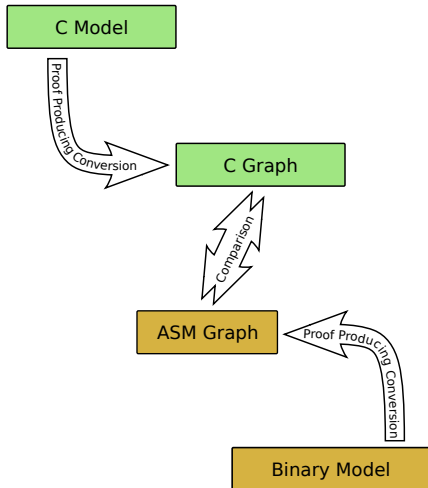
Approach



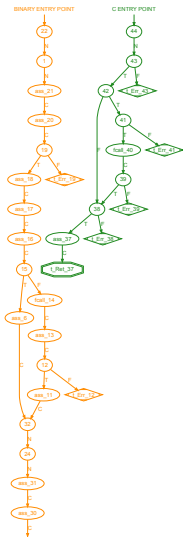
Approach



NICTA

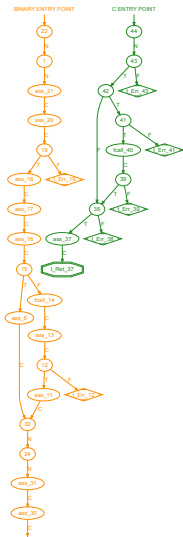


Proving Graph Refinement



The big challenge is the inner graph refinement.
This is proven one function at a time.

Proving Graph Refinement

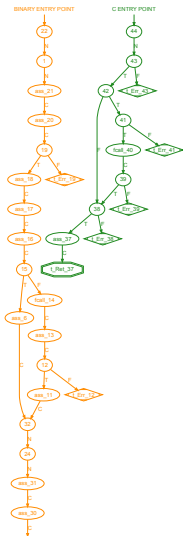


The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms.

Proving Graph Refinement

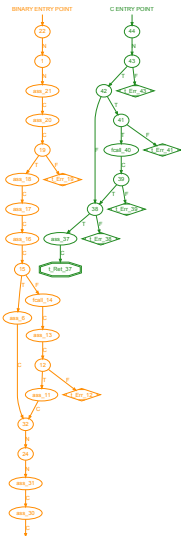


The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms. ❌

Proving Graph Refinement

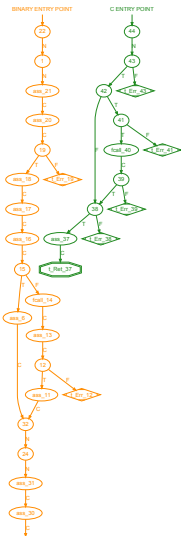


The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms. ❌
- Showing equivalences one basic block at a time.

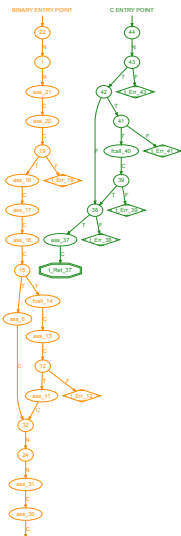
Proving Graph Refinement



The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms. ❌
- Showing equivalences one basic block at a time. ❌

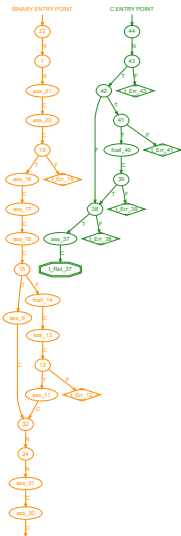


The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms. ❌
- Showing equivalences one basic block at a time. ❌
- Conversion of whole problems to SMT

Proving Graph Refinement



The big challenge is the inner graph refinement.
This is proven one function at a time.

Proven by:

- Implementing compiler-like transforms. ❌
- Showing equivalences one basic block at a time. ❌
- Conversion of whole problems to SMT
✓ modulo cycles.

What about cycles?

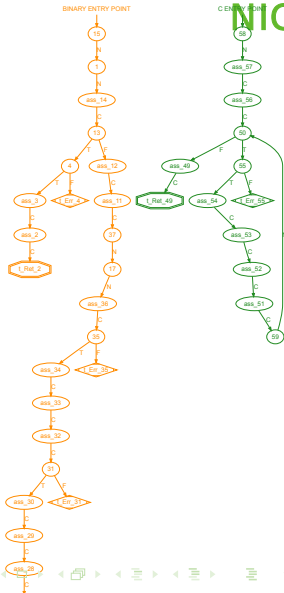


NICTA

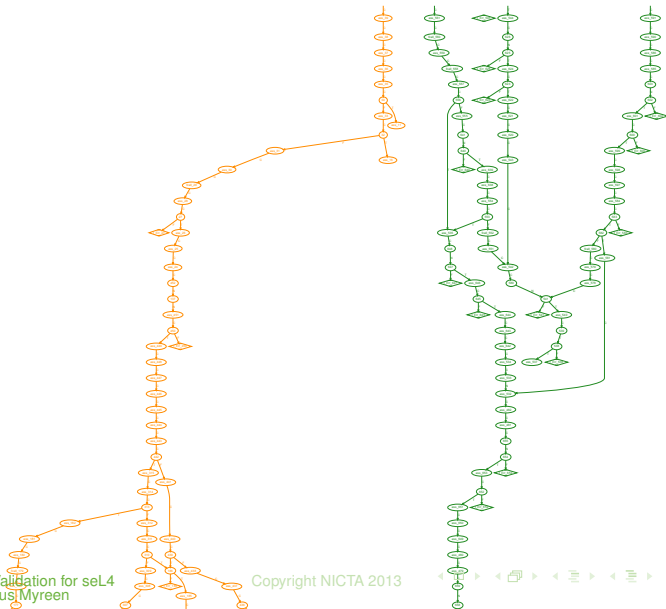
What about cycles?

We have two approaches:

- 1 Discover a loop bound.
- 2 Perform split point induction.



Challenge 1



Challenges:

- Inlining & problem size.
- Counterexample size.
- Finding split induction parameters.
- Functions marked `const` or `pure`.
- Partiality from C standard, binary semantics, decompiler.
- SMT theory extension for C standard symbols.
- Special memory regions:
 - Pointer memory regions (types matter for strict-aliasing).
 - Global objects.
 - ELF sections `.rodata` `.text` etc.
 - Usable Memory.

Results: Works for previously-verified seL4 code with gcc-4.5.1 -O1.
Nested loops and higher optimisation levels not yet handled.

Results: Works for previously-verified seL4 code with gcc-4.5.1 -O1.

Nested loops and higher optimisation levels not yet handled.

Conclusion: It is possible to build a certified compilation environment out of gcc, SMT and tape.