Veriopt

March 17, 2023

Abstract

The Veriopt project aims to prove the optimization pass of the GraalVM compiler. The GraalVM compiler includes a sophisticated Intermediate Representation (IR) in the form of a sea-of-nodes based graph structure. We first define the IR graph structure in the Isabelle/HOL interactive theorem prover. We subsequently give the evaluation of the structure a semantics based on the current understanding of the purpose of each IR graph node. Optimization phases are then encoded including the static analysis passes required for an optimization. Each optimization phase is proved to be correct by proving that a bisimulation exists between the unoptimized and optimized graphs. The following document has been automatically generated from the Isabelle/HOL source to provide a very comprehensive definition of the semantics and optimizations introduced by the Veriopt project.

Contents

1 Operator Semantics

theory *Values*
  imports
    HOL−Library.Word
    HOL−Library.Signed-Division
    HOL−Library.Float
    HOL−Library.LaTeXsugar
begin

In order to properly implement the IR semantics we first introduce a type that represents runtime values. These runtime values represent the full range of primitive types currently allowed by our semantics, ranging from basic integer types to object references and arrays.

Note that Java supports 64, 32, 16, 8 signed ints, plus 1 bit (boolean) ints, and char is 16-bit unsigned. E.g. an 8-bit stamp has a default range of -128..+127. And a 1-bit stamp has a default range of -1..0, surprisingly.

During calculations the smaller sizes are sign-extended to 32 bits, but explicit widening nodes will do that, so most binary calculations should see equal input sizes.

An object reference is an option type where the *None* object reference points to the static fields. This is examined more closely in our definition of the heap.

type-synonym *int64* = *64 word* — long
type-synonym *int32* = *32 word* — int
type-synonym *int16* = *16 word* — short
type-synonym *int8* = *8 word* — char
type-synonym *int1* = *1 word* — boolean

abbreviation *valid-int-widths* :: *nat set* **where**
  valid-int-widths ≡ {*1*, *8*, *16*, *32*, *64*}

experiment begin

Option 2: explicit width stored with each integer value. However, this does not help us to distinguish between short (signed) and char (unsigned).

typedef *IntWidth* = { *w* :: *nat* . *w=1* ∨ *w=8* ∨ *w=16* ∨ *w=32* ∨ *w=64* }
  ⟨*proof*⟩

setup-lifting *type-definition-IntWidth*

lift-definition *IntWidthBits* :: *IntWidth* ⇒ *nat*
  is λ*w*. *w* ⟨*proof*⟩
end

experiment begin

Option 3: explicit type stored with each integer value.

datatype *IntType* = *ILong* | *IInt* | *IShort* | *IChar* | *IByte* | *IBoolean*

fun *int-bits* :: *IntType* ⇒ *nat* **where**
  int-bits ILong = *64* |
  int-bits IInt = *32* |
  int-bits IShort = *16* |
  int-bits IChar = *16* |
  int-bits IByte = *8* |
  int-bits IBoolean = *1*

fun *int-signed* :: *IntType* ⇒ *bool* **where**
  int-signed ILong = *True* |
  int-signed IInt = *True* |
  int-signed IShort = *True* |
  int-signed IChar = *False* |
  int-signed IByte = *True* |
  int-signed IBoolean = *True*
end

Option 4: int64 with the number of significant bits.

type-synonym *iwidth* = *nat*
type-synonym *objref* = *nat option*

datatype (*discs-sels*) *Value* =
  UndefVal |

  IntVal iwidth int64 |

  ObjRef objref |
  ObjStr string

fun *intval-bits* :: *Value* ⇒ *nat* **where**
  intval-bits (*IntVal b v*) = *b*

fun *intval-word* :: *Value* ⇒ *int64* **where**
  intval-word (*IntVal b v*) = *v*

fun *bit-bounds* :: *nat* ⇒ (*int* × *int*) **where**
  bit-bounds bits = (((*2* ^ *bits*) *div 2*) ∗ −*1*, ((*2* ^ *bits*) *div 2*) − *1*)

definition *logic-negate* :: (′*a*::*len*) *word* ⇒ ′*a word* **where**
  logic-negate x = (*if x* = *0 then 1 else 0*)

fun *int-signed-value* :: *iwidth* ⇒ *int64* ⇒ *int* **where**
int-signed-value b v = *sint* (*signed-take-bit* (*b* − *1*) *v*)

fun *int-unsigned-value* :: *iwidth* ⇒ *int64* ⇒ *int* **where**
int-unsigned-value b v = *uint v*

Converts an integer word into a Java value.

fun *new-int* :: *iwidth* ⇒ *int64* ⇒ *Value* **where**
new-int b w = *IntVal b* (*take-bit b w*)

Converts an integer word into a Java value, iff the two types are equal.

fun *new-int-bin* :: *iwidth* ⇒ *iwidth* ⇒ *int64* ⇒ *Value* **where**
new-int-bin b1 b2 w = (*if b1=b2 then new-int b1 w else UndefVal*)

fun *wf-bool* :: *Value* ⇒ *bool* **where**
wf-bool (*IntVal b w*) = (*b* = *1*) |
wf-bool - = *False*

fun *val-to-bool* :: *Value* ⇒ *bool* **where**
val-to-bool (*IntVal b val*) = (*if val* = *0 then False else True*) |
val-to-bool val = *False*

fun *bool-to-val* :: *bool* ⇒ *Value* **where**
bool-to-val True = (*IntVal 32 1*) |
bool-to-val False = (*IntVal 32 0*)

Converts an Isabelle bool into a Java value, iff the two types are equal.

fun *bool-to-val-bin* :: *iwidth* ⇒ *iwidth* ⇒ *bool* ⇒ *Value* **where**
bool-to-val-bin t1 t2 b = (*if t1* = *t2 then bool-to-val b else UndefVal*)

fun *is-int-val* :: *Value* ⇒ *bool* **where**
is-int-val v = *is-IntVal v*

A convenience function for directly constructing -1 values of a given bit size.

fun *neg-one* :: *iwidth* ⇒ *int64* **where**
neg-one b = *mask b*

lemma *neg-one-value*[*simp*]: *new-int b* (*neg-one b*) = *IntVal b* (*mask b*)
⟨*proof*⟩

lemma *neg-one-signed*[*simp*]:
assumes *0* < *b*
shows *int-signed-value b* (*neg-one b*) = −*1*
⟨*proof*⟩

1.1 Arithmetic Operators

We need to introduce arithmetic operations which agree with the JVM.

Within the JVM, bytecode arithmetic operations are performed on 32 or 64 bit integers, unboxing where appropriate.

The following collection of intval functions correspond to the JVM arithmetic operations. We merge the 32 and 64 bit operations into a single function, even though the stamp of each IRNode tells us exactly what the bit widths will be. These merged functions make it easier to do the instantiation of Value as 'plus', etc. It might be worse for reasoning, because it could cause more case analysis, but this does not seem to be a problem in practice.

fun *intval-add* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-add (*IntVal b1 v1*) (*IntVal b2 v2*) =
(*if b1* = *b2 then IntVal b1* (*take-bit b1* (*v1*+*v2*)) *else UndefVal*) |
intval-add - - = *UndefVal*

fun *intval-sub* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-sub (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int-bin b1 b2* (*v1*−*v2*) |
intval-sub - - = *UndefVal*

fun *intval-mul* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-mul (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int-bin b1 b2* (*v1*∗*v2*) |
intval-mul - - = *UndefVal*

fun *intval-div* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-div (*IntVal b1 v1*) (*IntVal b2 v2*) =
new-int-bin b1 b2 (*word-of-int*
((*int-signed-value b1 v1*) *sdiv* (*int-signed-value b2 v2*))) |
intval-div - - = *UndefVal*

fun *intval-mod* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-mod (*IntVal b1 v1*) (*IntVal b2 v2*) =
new-int-bin b1 b2 (*word-of-int*
((*int-signed-value b1 v1*) *smod* (*int-signed-value b2 v2*))) |
intval-mod - - = *UndefVal*

fun *intval-negate* :: *Value* ⇒ *Value* **where**
intval-negate (*IntVal t v*) = *new-int t* (− *v*) |
intval-negate - = *UndefVal*

fun *intval-abs* :: *Value* ⇒ *Value* **where**

intval-abs (*IntVal t v*) = *new-int t* (*if int-signed-value t v* < *0 then* − *v else v*) |
intval-abs - = *UndefVal*

TODO: clarify which widths this should work on: just 1-bit or all?

fun *intval-logic-negation* :: *Value* ⇒ *Value* **where**
intval-logic-negation (*IntVal b v*) = *new-int b* (*logic-negate v*) |
intval-logic-negation - = *UndefVal*

1.2 Bitwise Operators

fun *intval-and* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-and (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int-bin b1 b2* (*and v1 v2*) |
intval-and - - = *UndefVal*

fun *intval-or* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-or (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int-bin b1 b2* (*or v1 v2*) |
intval-or - - = *UndefVal*

fun *intval-xor* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-xor (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int-bin b1 b2* (*xor v1 v2*) |
intval-xor - - = *UndefVal*

fun *intval-not* :: *Value* ⇒ *Value* **where**
intval-not (*IntVal t v*) = *new-int t* (*not v*) |
intval-not - = *UndefVal*

1.3 Comparison Operators

fun *intval-short-circuit-or* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-short-circuit-or (*IntVal b1 v1*) (*IntVal b2 v2*) = *bool-to-val-bin b1 b2* (((*v1* ≠ *0*) ∨ (*v2* ≠ *0*))) |
intval-short-circuit-or - - = *UndefVal*

fun *intval-equals* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-equals (*IntVal b1 v1*) (*IntVal b2 v2*) = *bool-to-val-bin b1 b2* (*v1* = *v2*) |
intval-equals - - = *UndefVal*

fun *intval-less-than* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-less-than (*IntVal b1 v1*) (*IntVal b2 v2*) =
bool-to-val-bin b1 b2 (*int-signed-value b1 v1* < *int-signed-value b2 v2*) |
intval-less-than - - = *UndefVal*

fun *intval-below* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-below (*IntVal b1 v1*) (*IntVal b2 v2*) = *bool-to-val-bin b1 b2* (*v1* < *v2*) |
intval-below - - = *UndefVal*

fun *intval-conditional* :: *Value* ⇒ *Value* ⇒ *Value* ⇒ *Value* **where**
intval-conditional cond tv fv = (*if* (*val-to-bool cond*) *then tv else fv*)

1.4 Narrowing and Widening Operators

Note: we allow these operators to have inBits=outBits, because the Graal compiler also seems to allow that case, even though it should rarely / never arise in practice.

value *sint*(*signed-take-bit 0* (*1* :: *int32*))

fun *intval-narrow* :: *nat* ⇒ *nat* ⇒ *Value* ⇒ *Value* **where**
  intval-narrow inBits outBits (*IntVal b v*) =
    (*if inBits* = *b* ∧ *0* < *outBits* ∧ *outBits* ≤ *inBits* ∧ *inBits* ≤ *64*
     then new-int outBits v
     else UndefVal) |
  intval-narrow - - - = *UndefVal*

value *sint* (*signed-take-bit 7* ((*256* + *128*) :: *int64*))

fun *intval-sign-extend* :: *nat* ⇒ *nat* ⇒ *Value* ⇒ *Value* **where**
  intval-sign-extend inBits outBits (*IntVal b v*) =
    (*if inBits* = *b* ∧ *0* < *inBits* ∧ *inBits* ≤ *outBits* ∧ *outBits* ≤ *64*
     then new-int outBits (*signed-take-bit* (*inBits* − *1*) *v*)
     else UndefVal) |
  intval-sign-extend - - - = *UndefVal*

fun *intval-zero-extend* :: *nat* ⇒ *nat* ⇒ *Value* ⇒ *Value* **where**
  intval-zero-extend inBits outBits (*IntVal b v*) =
    (*if inBits* = *b* ∧ *0* < *inBits* ∧ *inBits* ≤ *outBits* ∧ *outBits* ≤ *64*
     then new-int outBits (*take-bit inBits v*)
     else UndefVal) |
  intval-zero-extend - - - = *UndefVal*

Some well-formedness results to help reasoning about narrowing and widening operators

lemma *intval-narrow-ok*:
  assumes *intval-narrow inBits outBits val* ≠ *UndefVal*
  shows *0* < *outBits* ∧ *outBits* ≤ *inBits* ∧ *inBits* ≤ *64* ∧ *outBits* ≤ *64* ∧
      is-IntVal val ∧
      intval-bits val = *inBits*
  ⟨*proof*⟩

lemma *intval-sign-extend-ok*:
  assumes *intval-sign-extend inBits outBits val* ≠ *UndefVal*
  shows *0* < *inBits* ∧
      inBits ≤ *outBits* ∧ *outBits* ≤ *64* ∧
      is-IntVal val ∧
      intval-bits val = *inBits*
  ⟨*proof*⟩

lemma *intval-zero-extend-ok*:
  assumes *intval-zero-extend inBits outBits val* ≠ *UndefVal*

shows $0 < inBits \wedge$
$inBits \leq outBits \wedge outBits \leq 64 \wedge$
$is\text{-}IntVal\ val \wedge$
$intval\text{-}bits\ val = inBits$
⟨proof⟩

1.5 Bit-Shifting Operators

definition *shiftl* (**infix** << *75*) **where**
shiftl w n = (*push-bit n*) *w*

lemma *shiftl-power*[*simp*]: $(x::('a::len)\ word) * (2\ \hat{}\ j) = x << j$
⟨proof⟩

lemma $(x::('a::len)\ word) * ((2\ \hat{}\ j) + 1) = x << j + x$
⟨proof⟩

lemma $(x::('a::len)\ word) * ((2\ \hat{}\ j) - 1) = x << j - x$
⟨proof⟩

lemma $(x::('a::len)\ word) * ((2\hat{}j) + (2\hat{}k)) = x << j + x << k$
⟨proof⟩

lemma $(x::('a::len)\ word) * ((2\hat{}j) - (2\hat{}k)) = x << j - x << k$
⟨proof⟩

definition *shiftr* (**infix** >>> *75*) **where**
shiftr w n = (*drop-bit n*) *w*

value (*255* :: *8 word*) >>> (*2* :: *nat*)

definition *sshiftr* :: *'a* :: *len word* ⇒ *nat* ⇒ *'a* :: *len word* (**infix** >> *75*) **where**
sshiftr w n = *word-of-int* ((*sint w*) *div* (*2* ^ *n*))

value (*128* :: *8 word*) >> *2*

Note that Java shift operators use unary numeric promotion, unlike other binary operators, which use binary numeric promotion (see the Java language reference manual). This means that the left-hand input determines the output size, while the right-hand input can be any size.

fun *shift-amount* :: *iwidth* ⇒ *int64* ⇒ *nat* **where**
shift-amount b val = *unat* (*and val* (*if b* = *64 then 0x3F else 0x1f*))

fun *intval-left-shift* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-left-shift (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int b1* (*v1* << *shift-amount b1 v2*) |
intval-left-shift - - = *UndefVal*

Signed shift is more complex, because we sometimes have to insert 1 bits at the correct point, which is at b1 bits.

fun *intval-right-shift* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-right-shift (*IntVal b1 v1*) (*IntVal b2 v2*) =
(*let shift* = *shift-amount b1 v2 in*
let ones = *and* (*mask b1*) (*not* (*mask* (*b1* − *shift*) :: *int64*)) *in*
(*if int-signed-value b1 v1* < *0*
then new-int b1 (*or ones* (*v1* >>> *shift*))
else new-int b1 (*v1* >>> *shift*))) |
intval-right-shift - - = *UndefVal*

fun *intval-uright-shift* :: *Value* ⇒ *Value* ⇒ *Value* **where**
intval-uright-shift (*IntVal b1 v1*) (*IntVal b2 v2*) = *new-int b1* (*v1* >>> *shift-amount b1 v2*) |
intval-uright-shift - - = *UndefVal*

1.5.1 Examples of Narrowing / Widening Functions

experiment begin
corollary *intval-narrow 32 8* (*IntVal 32* (*256* + *128*)) = *IntVal 8 128* ⟨*proof*⟩
corollary *intval-narrow 32 8* (*IntVal 32* (−*2*)) = *IntVal 8 254* ⟨*proof*⟩
corollary *intval-narrow 32 1* (*IntVal 32* (−*2*)) = *IntVal 1 0* ⟨*proof*⟩
corollary *intval-narrow 32 1* (*IntVal 32* (−*3*)) = *IntVal 1 1* ⟨*proof*⟩

corollary *intval-narrow 32 8* (*IntVal 64* (−*2*)) = *UndefVal* ⟨*proof*⟩
corollary *intval-narrow 64 8* (*IntVal 32* (−*2*)) = *UndefVal* ⟨*proof*⟩
corollary *intval-narrow 64 8* (*IntVal 64 254*) = *IntVal 8 254* ⟨*proof*⟩
corollary *intval-narrow 64 8* (*IntVal 64* (*256*+*127*)) = *IntVal 8 127* ⟨*proof*⟩
corollary *intval-narrow 64 64* (*IntVal 64* (−*2*)) = *IntVal 64* (−*2*) ⟨*proof*⟩
end

experiment begin
corollary *intval-sign-extend 8 32* (*IntVal 8* (*256* + *128*)) = *IntVal 32* (*2*^*32* − *128*) ⟨*proof*⟩
corollary *intval-sign-extend 8 32* (*IntVal 8* (−*2*)) = *IntVal 32* (*2*^*32* − *2*) ⟨*proof*⟩
corollary *intval-sign-extend 1 32* (*IntVal 1* (−*2*)) = *IntVal 32 0* ⟨*proof*⟩
corollary *intval-sign-extend 1 32* (*IntVal 1* (−*3*)) = *IntVal 32* (*mask 32*) ⟨*proof*⟩

corollary *intval-sign-extend 8 32* (*IntVal 64 254*) = *UndefVal* ⟨*proof*⟩
corollary *intval-sign-extend 8 64* (*IntVal 32 254*) = *UndefVal* ⟨*proof*⟩
corollary *intval-sign-extend 8 64* (*IntVal 8 254*) = *IntVal 64* (−*2*) ⟨*proof*⟩
corollary *intval-sign-extend 32 64* (*IntVal 32* (*2*^*32* − *2*)) = *IntVal 64* (−*2*) ⟨*proof*⟩
corollary *intval-sign-extend 64 64* (*IntVal 64* (−*2*)) = *IntVal 64* (−*2*) ⟨*proof*⟩
end

experiment begin

corollary *intval-zero-extend 8 32* (*IntVal 8* (*256* + *128*)) = *IntVal 32 128* ⟨*proof*⟩
corollary *intval-zero-extend 8 32* (*IntVal 8* (−*2*)) = *IntVal 32 254* ⟨*proof*⟩
corollary *intval-zero-extend 1 32* (*IntVal 1* (−*1*)) = *IntVal 32 1* ⟨*proof*⟩
corollary *intval-zero-extend 1 32* (*IntVal 1* (−*2*)) = *IntVal 32 0* ⟨*proof*⟩

corollary *intval-zero-extend 8 32* (*IntVal 64* (−*2*)) = *UndefVal* ⟨*proof*⟩
corollary *intval-zero-extend 8 64* (*IntVal 64* (−*2*)) = *UndefVal* ⟨*proof*⟩
corollary *intval-zero-extend 8 64* (*IntVal 8 254*) = *IntVal 64 254* ⟨*proof*⟩
corollary *intval-zero-extend 32 64* (*IntVal 32* (*2*^*32* − *2*)) = *IntVal 64* (*2*^*32* − *2*) ⟨*proof*⟩
corollary *intval-zero-extend 64 64* (*IntVal 64* (−*2*)) = *IntVal 64* (−*2*) ⟨*proof*⟩
end

experiment begin
corollary *intval-right-shift* (*IntVal 8 128*) (*IntVal 8 0*) = *IntVal 8 128* ⟨*proof*⟩
corollary *intval-right-shift* (*IntVal 8 128*) (*IntVal 8 1*) = *IntVal 8 192* ⟨*proof*⟩
corollary *intval-right-shift* (*IntVal 8 128*) (*IntVal 8 2*) = *IntVal 8 224* ⟨*proof*⟩
corollary *intval-right-shift* (*IntVal 8 128*) (*IntVal 8 8*) = *IntVal 8 255* ⟨*proof*⟩
corollary *intval-right-shift* (*IntVal 8 128*) (*IntVal 8 31*) = *IntVal 8 255* ⟨*proof*⟩
end

lemma *intval-add-sym*:
shows *intval-add a b* = *intval-add b a*
⟨*proof*⟩

code-deps *intval-add*
code-thms *intval-add*

lemma *intval-add* (*IntVal 32* (*2*^*31*−*1*)) (*IntVal 32* (*2*^*31*−*1*)) = *IntVal 32* (*2*^*32* − *2*)
⟨*proof*⟩
lemma *intval-add* (*IntVal 64* (*2*^*31*−*1*)) (*IntVal 64* (*2*^*31*−*1*)) = *IntVal 64 4294967294*
⟨*proof*⟩

end

1.6 Fixed-width Word Theories

theory *ValueThms*
imports *Values*
begin

1.6.1 Support Lemmas for Upper/Lower Bounds

lemma *size32*: *size v* = *32* **for** *v* :: *32 word*
⟨*proof*⟩

lemma *size64*: *size v* = *64* **for** *v* :: *64 word*
⟨*proof*⟩

lemma *lower-bounds-equiv*:
assumes $0 < N$
shows $-(((2::int) \ \hat{} \ (N-1))) = (2::int) \ \hat{} \ N \ div \ 2 * - 1$
⟨*proof*⟩

lemma *upper-bounds-equiv*:
assumes $0 < N$
shows $(2::int) \ \hat{} \ (N-1) = (2::int) \ \hat{} \ N \ div \ 2$
⟨*proof*⟩

Some min/max bounds for 64-bit words

lemma *bit-bounds-min64*: $((fst\ (bit\text{-}bounds\ 64))) \leq (sint\ (v::int64))$
⟨*proof*⟩

lemma *bit-bounds-max64*: $((snd\ (bit\text{-}bounds\ 64))) \geq (sint\ (v::int64))$
⟨*proof*⟩

Extend these min/max bounds to extracting smaller signed words using *signed_take_bit*.

Note: we could use signed to convert between bit-widths, instead of signed_take_bit. But that would have to be done separately for each bit-width type.

value *sint*(*signed-take-bit 7* (*128* :: *int8*))

ML-val ‹@{*thm signed-take-bit-decr-length-iff*}›
declare [[*show-types*=*true*]]
ML-val ‹@{*thm signed-take-bit-int-less-exp*}›

lemma *signed-take-bit-int-less-exp-word*:
fixes *ival* :: *'a* :: *len word*
assumes $n < LENGTH('a)$
shows $sint(signed\text{-}take\text{-}bit\ n\ ival) < (2::int) \ \hat{} \ n$
⟨*proof*⟩

lemma *signed-take-bit-int-greater-eq-minus-exp-word*:
fixes *ival* :: *'a* :: *len word*
assumes $n < LENGTH('a)$

shows $-(2 \mathbin{\hat{}} n) \le sint(signed\text{-}take\text{-}bit\ n\ ival)$
⟨*proof*⟩

lemma *signed-take-bit-range*:
fixes $ival :: 'a :: len\ word$
assumes $n < LENGTH('a)$
assumes $val = sint(signed\text{-}take\text{-}bit\ n\ ival)$
shows $-(2 \mathbin{\hat{}} n) \le val \wedge val < 2 \mathbin{\hat{}} n$
⟨*proof*⟩

A *bit_bounds* version of the above lemma.

lemma *signed-take-bit-bounds*:
fixes $ival :: 'a :: len\ word$
assumes $n \le LENGTH('a)$
assumes $0 < n$
assumes $val = sint(signed\text{-}take\text{-}bit\ (n - 1)\ ival)$
shows $fst\ (bit\text{-}bounds\ n) \le val \wedge val \le snd\ (bit\text{-}bounds\ n)$
⟨*proof*⟩

lemma *signed-take-bit-bounds64*:
fixes $ival :: int64$
assumes $n \le 64$
assumes $0 < n$
assumes $val = sint(signed\text{-}take\text{-}bit\ (n - 1)\ ival)$
shows $fst\ (bit\text{-}bounds\ n) \le val \wedge val \le snd\ (bit\text{-}bounds\ n)$
⟨*proof*⟩

lemma *int-signed-value-bounds*:
assumes $b1 \le 64$
assumes $0 < b1$
shows $fst\ (bit\text{-}bounds\ b1) \le int\text{-}signed\text{-}value\ b1\ v2 \wedge$
$int\text{-}signed\text{-}value\ b1\ v2 \le snd\ (bit\text{-}bounds\ b1)$
⟨*proof*⟩

lemma *int-signed-value-range*:
fixes $ival :: int64$
assumes $val = int\text{-}signed\text{-}value\ n\ ival$
shows $-(2 \mathbin{\hat{}} (n - 1)) \le val \wedge val < 2 \mathbin{\hat{}} (n - 1)$
⟨*proof*⟩

Some lemmas about unsigned words smaller than 64-bit, for zero-extend operators.

lemma *take-bit-smaller-range*:
fixes $ival :: 'a :: len\ word$
assumes $n < LENGTH('a)$
assumes $val = sint(take\text{-}bit\ n\ ival)$
shows $0 \le val \wedge val < (2::int) \mathbin{\hat{}} n$
⟨*proof*⟩

lemma *take-bit-same-size-nochange*:
  fixes *ival* :: *'a* :: *len word*
  assumes $n = LENGTH('a)$
  shows *ival* = *take-bit n ival*
  ⟨*proof*⟩

A simplification lemma for *new_int*, showing that upper bits can be ignored.

lemma *take-bit-redundant*[*simp*]:
  fixes *ival* :: *'a* :: *len word*
  assumes $0 < n$
  assumes $n < LENGTH('a)$
  shows *signed-take-bit* (*n* − *1*) (*take-bit n ival*) = *signed-take-bit* (*n* − *1*) *ival*
⟨*proof*⟩

lemma *take-bit-same-size-range*:
  fixes *ival* :: *'a* :: *len word*
  assumes $n = LENGTH('a)$
  assumes *ival2* = *take-bit n ival*
  shows − (*2* ^ *n div 2*) ≤ *sint ival2* ∧ *sint ival2* < *2* ^ *n div 2*
  ⟨*proof*⟩

lemma *take-bit-same-bounds*:
  fixes *ival* :: *'a* :: *len word*
  assumes $n = LENGTH('a)$
  assumes *ival2* = *take-bit n ival*
  shows *fst* (*bit-bounds n*) ≤ *sint ival2* ∧ *sint ival2* ≤ *snd* (*bit-bounds n*)
  ⟨*proof*⟩

Next we show that casting a word to a wider word preserves any upper/lower bounds. (These lemmas may not be needed any more, since we are not using scast now?)

lemma *scast-max-bound*:
  assumes *sint* (*v* :: *'a* :: *len word*) < *M*
  assumes $LENGTH('a) < LENGTH('b)$
  shows *sint* ((*scast v*) :: *'b* :: *len word*) < *M*
  ⟨*proof*⟩

lemma *scast-min-bound*:
  assumes *M* ≤ *sint* (*v* :: *'a* :: *len word*)
  assumes $LENGTH('a) < LENGTH('b)$
  shows *M* ≤ *sint* ((*scast v*) :: *'b* :: *len word*)
  ⟨*proof*⟩

lemma *scast-bigger-max-bound*:
  assumes (*result* :: *'b* :: *len word*) = *scast* (*v* :: *'a* :: *len word*)
  shows *sint result* < *2* ^ *LENGTH*(*'a*) *div 2*
  ⟨*proof*⟩

lemma *scast-bigger-min-bound*:
assumes $(result :: 'b :: len\ word) = scast\ (v :: 'a :: len\ word)$
shows $-\ (2\ \hat{}\ LENGTH('a)\ div\ 2) \leq sint\ result$
⟨*proof*⟩

lemma *scast-bigger-bit-bounds*:
assumes $(result :: 'b :: len\ word) = scast\ (v :: 'a :: len\ word)$
shows $fst\ (bit\text{-}bounds\ (LENGTH('a))) \leq sint\ result \wedge sint\ result \leq snd\ (bit\text{-}bounds\ (LENGTH('a)))$
⟨*proof*⟩

Results about *new_int*.

lemma *new-int-take-bits*:
assumes $IntVal\ b\ val = new\text{-}int\ b\ ival$
shows $take\text{-}bit\ b\ val = val$
⟨*proof*⟩

1.6.2 Support lemmas for take bit and signed take bit.

Lemmas for removing redundant take_bit wrappers.

lemma *take-bit-dist-addL*[*simp*]:
fixes $x :: 'a :: len\ word$
shows $take\text{-}bit\ b\ (take\text{-}bit\ b\ x + y) = take\text{-}bit\ b\ (x + y)$
⟨*proof*⟩

lemma *take-bit-dist-addR*[*simp*]:
fixes $x :: 'a :: len\ word$
shows $take\text{-}bit\ b\ (x + take\text{-}bit\ b\ y) = take\text{-}bit\ b\ (x + y)$
⟨*proof*⟩

lemma *take-bit-dist-subL*[*simp*]:
fixes $x :: 'a :: len\ word$
shows $take\text{-}bit\ b\ (take\text{-}bit\ b\ x - y) = take\text{-}bit\ b\ (x - y)$
⟨*proof*⟩

lemma *take-bit-dist-subR*[*simp*]:
fixes $x :: 'a :: len\ word$
shows $take\text{-}bit\ b\ (x - take\text{-}bit\ b\ y) = take\text{-}bit\ b\ (x - y)$
⟨*proof*⟩

lemma *take-bit-dist-neg*[*simp*]:
fixes $ix :: 'a :: len\ word$
shows $take\text{-}bit\ b\ (-\ take\text{-}bit\ b\ (ix)) = take\text{-}bit\ b\ (-\ ix)$
⟨*proof*⟩

lemma *signed-take-take-bit*[*simp*]:
  fixes $x :: 'a :: len\ word$
  assumes $0 < b$
  shows *signed-take-bit* $(b - 1)$ (*take-bit* $b\ x$) $=$ *signed-take-bit* $(b - 1)\ x$
  ⟨*proof*⟩

lemma *mod-larger-ignore*:
  fixes $a :: int$
  fixes $m\ n :: nat$
  assumes $n < m$
  shows $(a\ mod\ 2\ \hat{}\ m)\ mod\ 2\ \hat{}\ n = a\ mod\ 2\ \hat{}\ n$
  ⟨*proof*⟩

lemma *mod-dist-over-add*:
  fixes $a\ b\ c :: int64$
  fixes $n :: nat$
  assumes *1*: $0 < n$
  assumes *2*: $n < 64$
  shows $(a\ mod\ 2\hat{}n + b)\ mod\ 2\hat{}n = (a + b)\ mod\ 2\hat{}n$
⟨*proof*⟩

end

2 Stamp Typing

theory *Stamp*
  imports *Values*
begin

The GraalVM compiler uses the Stamp class to store range and type information for a given node in the IR graph. We model the Stamp class as a datatype, Stamp, and provide a number of functions on the datatype which correspond to the class methods within the compiler.

Stamp information is used in a variety of ways in optimizations, and so, we additionally provide a number of lemmas which help to prove future optimizations.

datatype *Stamp* =
  VoidStamp
  | *IntegerStamp* (*stp-bits*: *nat*) (*stpi-lower*: *int*) (*stpi-upper*: *int*)

  | *KlassPointerStamp* (*stp-nonNull*: *bool*) (*stp-alwaysNull*: *bool*)
  | *MethodCountersPointerStamp* (*stp-nonNull*: *bool*) (*stp-alwaysNull*: *bool*)
  | *MethodPointersStamp* (*stp-nonNull*: *bool*) (*stp-alwaysNull*: *bool*)
 | *ObjectStamp* (*stp-type*: *string*) (*stp-exactType*: *bool*) (*stp-nonNull*: *bool*) (*stp-alwaysNull*: *bool*)
  | *RawPointerStamp* (*stp-nonNull*: *bool*) (*stp-alwaysNull*: *bool*)

| *IllegalStamp*

fun *is-stamp-empty* :: *Stamp* ⇒ *bool* **where**
is-stamp-empty (*IntegerStamp b lower upper*) = (*upper* < *lower*) |

is-stamp-empty x = *False*

Just like the IntegerStamp class, we need to know that our lo/hi bounds fit into the given number of bits (either signed or unsigned). Our integer stamps have infinite lo/hi bounds, so if the lower bound is non-negative, we can assume that all values are positive, and the integer bits of a related value can be interpreted as unsigned. This is similar (but slightly more general) to what IntegerStamp.java does with its test: if (sameSignBounds()) in the unsignedUpperBound() method.

Note that this is a bit different and more accurate than what StampFactory.forUnsignedInteger does (it widens large unsigned ranges to the max signed range to allow all bit patterns) because its lo/hi values are only 64-bit.

fun *valid-stamp* :: *Stamp* ⇒ *bool* **where**
valid-stamp (*IntegerStamp bits lo hi*) =
(*0* < *bits* ∧ *bits* ≤ *64* ∧
fst (*bit-bounds bits*) ≤ *lo* ∧ *lo* ≤ *snd* (*bit-bounds bits*) ∧
fst (*bit-bounds bits*) ≤ *hi* ∧ *hi* ≤ *snd* (*bit-bounds bits*)) |
valid-stamp s = *True*

experiment begin
corollary *bit-bounds 1* = (−*1*, *0*) ⟨*proof*⟩
end

— A stamp which includes the full range of the type
fun *unrestricted-stamp* :: *Stamp* ⇒ *Stamp* **where**
unrestricted-stamp VoidStamp = *VoidStamp* |
unrestricted-stamp (*IntegerStamp bits lower upper*) = (*IntegerStamp bits* (*fst* (*bit-bounds bits*)) (*snd* (*bit-bounds bits*))) |

unrestricted-stamp (*KlassPointerStamp nonNull alwaysNull*) = (*KlassPointerStamp False False*) |
unrestricted-stamp (*MethodCountersPointerStamp nonNull alwaysNull*) = (*MethodCountersPointerStamp False False*) |
unrestricted-stamp (*MethodPointersStamp nonNull alwaysNull*) = (*MethodPointersStamp*

False False) |
unrestricted-stamp (*ObjectStamp type exactType nonNull alwaysNull*) = (*ObjectStamp* '''' *False False False*) |
unrestricted-stamp - = *IllegalStamp*

fun *is-stamp-unrestricted* :: *Stamp* ⇒ *bool* **where**
is-stamp-unrestricted s = (*s* = *unrestricted-stamp s*)

— A stamp which provides type information but has an empty range of values
fun *empty-stamp* :: *Stamp* ⇒ *Stamp* **where**
empty-stamp VoidStamp = *VoidStamp* |
empty-stamp (*IntegerStamp bits lower upper*) = (*IntegerStamp bits* (*snd* (*bit-bounds bits*)) (*fst* (*bit-bounds bits*))) |

empty-stamp (*KlassPointerStamp nonNull alwaysNull*) = (*KlassPointerStamp nonNull alwaysNull*) |
empty-stamp (*MethodCountersPointerStamp nonNull alwaysNull*) = (*MethodCountersPointerStamp nonNull alwaysNull*) |
empty-stamp (*MethodPointersStamp nonNull alwaysNull*) = (*MethodPointersStamp nonNull alwaysNull*) |
empty-stamp (*ObjectStamp type exactType nonNull alwaysNull*) = (*ObjectStamp* '''' *True True False*) |
empty-stamp stamp = *IllegalStamp*

— Calculate the meet stamp of two stamps
fun *meet* :: *Stamp* ⇒ *Stamp* ⇒ *Stamp* **where**
meet VoidStamp VoidStamp = *VoidStamp* |
meet (*IntegerStamp b1 l1 u1*) (*IntegerStamp b2 l2 u2*) = (
if b1 ≠ *b2 then IllegalStamp else*
(*IntegerStamp b1* (*min l1 l2*) (*max u1 u2*))
) |

meet (*KlassPointerStamp nn1 an1*) (*KlassPointerStamp nn2 an2*) = (
KlassPointerStamp (*nn1* ∧ *nn2*) (*an1* ∧ *an2*)
) |
meet (*MethodCountersPointerStamp nn1 an1*) (*MethodCountersPointerStamp nn2 an2*) = (
MethodCountersPointerStamp (*nn1* ∧ *nn2*) (*an1* ∧ *an2*)
) |
meet (*MethodPointersStamp nn1 an1*) (*MethodPointersStamp nn2 an2*) = (
MethodPointersStamp (*nn1* ∧ *nn2*) (*an1* ∧ *an2*)
) |
meet s1 s2 = *IllegalStamp*

— Calculate the join stamp of two stamps
fun *join* :: *Stamp* ⇒ *Stamp* ⇒ *Stamp* **where**
join VoidStamp VoidStamp = *VoidStamp* |
join (*IntegerStamp b1 l1 u1*) (*IntegerStamp b2 l2 u2*) = (

$if\ b1 \neq b2\ then\ IllegalStamp\ else$
$(IntegerStamp\ b1\ (max\ l1\ l2)\ (min\ u1\ u2))$
$) \mid$

$join\ (KlassPointerStamp\ nn1\ an1)\ (KlassPointerStamp\ nn2\ an2) = ($
$if\ ((nn1 \vee nn2) \wedge (an1 \vee an2))$
$then\ (empty\text{-}stamp\ (KlassPointerStamp\ nn1\ an1))$
$else\ (KlassPointerStamp\ (nn1 \vee nn2)\ (an1 \vee an2))$
$) \mid$
$join\ (MethodCountersPointerStamp\ nn1\ an1)\ (MethodCountersPointerStamp\ nn2\ an2) = ($
$if\ ((nn1 \vee nn2) \wedge (an1 \vee an2))$
$then\ (empty\text{-}stamp\ (MethodCountersPointerStamp\ nn1\ an1))$
$else\ (MethodCountersPointerStamp\ (nn1 \vee nn2)\ (an1 \vee an2))$
$) \mid$
$join\ (MethodPointersStamp\ nn1\ an1)\ (MethodPointersStamp\ nn2\ an2) = ($
$if\ ((nn1 \vee nn2) \wedge (an1 \vee an2))$
$then\ (empty\text{-}stamp\ (MethodPointersStamp\ nn1\ an1))$
$else\ (MethodPointersStamp\ (nn1 \vee nn2)\ (an1 \vee an2))$
$) \mid$
$join\ s1\ s2 = IllegalStamp$

— In certain circumstances a stamp provides enough information to evaluate a value as a stamp, the asConstant function converts the stamp to a value where one can be inferred.

fun *asConstant* :: *Stamp* ⇒ *Value* **where**
$asConstant\ (IntegerStamp\ b\ l\ h) = (if\ l = h\ then\ IntVal\ b\ (word\text{-}of\text{-}int\ l)\ else\ UndefVal) \mid$
$asConstant\ \text{-} = UndefVal$

— Determine if two stamps never have value overlaps i.e. their join is empty
fun *alwaysDistinct* :: *Stamp* ⇒ *Stamp* ⇒ *bool* **where**
$alwaysDistinct\ stamp1\ stamp2 = is\text{-}stamp\text{-}empty\ (join\ stamp1\ stamp2)$

— Determine if two stamps must always be the same value i.e. two equal constants
fun *neverDistinct* :: *Stamp* ⇒ *Stamp* ⇒ *bool* **where**
$neverDistinct\ stamp1\ stamp2 = (asConstant\ stamp1 = asConstant\ stamp2 \wedge asConstant\ stamp1 \neq UndefVal)$

fun *constantAsStamp* :: *Value* ⇒ *Stamp* **where**
$constantAsStamp\ (IntVal\ b\ v) = (IntegerStamp\ b\ (int\text{-}signed\text{-}value\ b\ v)\ (int\text{-}signed\text{-}value\ b\ v)) \mid$

$constantAsStamp\ \text{-} = IllegalStamp$

— Define when a runtime value is valid for a stamp. The stamp bounds must be valid, and val must be zero-extended.
fun *valid-value* :: *Value* ⇒ *Stamp* ⇒ *bool* **where**

valid-value (*IntVal b1 val*) (*IntegerStamp b l h*) =
  (*if b1* = *b then*
    valid-stamp (*IntegerStamp b l h*) ∧
    take-bit b val = *val* ∧
    l ≤ *int-signed-value b val* ∧ *int-signed-value b val* ≤ *h*
  else False) |

valid-value (*ObjRef ref*) (*ObjectStamp klass exact nonNull alwaysNull*) =
  ((*alwaysNull* ⟶ *ref* = *None*) ∧ (*ref*=*None* ⟶ ¬ *nonNull*)) |
valid-value stamp val = *False*

definition *wf-value* :: *Value* ⇒ *bool* **where**
wf-value v = *valid-value v* (*constantAsStamp v*)

lemma *unfold-wf-value*[*simp*]:
wf-value v ⟹ *valid-value v* (*constantAsStamp v*)
⟨*proof*⟩

fun *compatible* :: *Stamp* ⇒ *Stamp* ⇒ *bool* **where**
compatible (*IntegerStamp b1 lo1 hi1*) (*IntegerStamp b2 lo2 hi2*) =
  (*b1* = *b2* ∧ *valid-stamp* (*IntegerStamp b1 lo1 hi1*) ∧ *valid-stamp* (*IntegerStamp b2 lo2 hi2*)) |
compatible (*VoidStamp*) (*VoidStamp*) = *True* |
compatible - - = *False*

fun *stamp-under* :: *Stamp* ⇒ *Stamp* ⇒ *bool* **where**
stamp-under (*IntegerStamp b1 lo1 hi1*) (*IntegerStamp b2 lo2 hi2*) = (*hi1* < *lo2*) |
stamp-under - - = *False*

— The most common type of stamp within the compiler (apart from the VoidStamp) is a 32 bit integer stamp with an unrestricted range. We use *default-stamp* as it is a frequently used stamp.
definition *default-stamp* :: *Stamp* **where**
default-stamp = (*unrestricted-stamp* (*IntegerStamp 32 0 0*))

value *valid-value* (*IntVal 8* (*255*)) (*IntegerStamp 8* (−*128*) *127*)
end

3 Graph Representation

3.1 IR Graph Nodes

theory *IRNodes*
imports
Values

begin

The GraalVM IR is represented using a graph data structure. Here we define the nodes that are contained within the graph. Each node represents a Node subclass in the GraalVM compiler, the node classes have annotated fields to indicate input and successor edges.

We represent these classes with each IRNode constructor explicitly labelling a reference to the node IDs that it stores as inputs and successors.

The inputs_of and successors_of functions partition those labelled references into input edges and successor edges of a node.

To identify each Node, we use a simple natural number index. Zero is always the start node in a graph. For human readability, within nodes we write INPUT (or special case thereof) instead of ID for input edges, and SUCC instead of ID for control-flow successor edges. Optional edges are handled as "INPUT option" etc.

type-synonym *ID = nat*
type-synonym *INPUT = ID*
type-synonym *INPUT-ASSOC = ID*
type-synonym *INPUT-STATE = ID*
type-synonym *INPUT-GUARD = ID*
type-synonym *INPUT-COND = ID*
type-synonym *INPUT-EXT = ID*
type-synonym *SUCC = ID*

datatype (*discs-sels*) *IRNode =*
AbsNode (*ir-value*: *INPUT*)
| *AddNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *AndNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *BeginNode* (*ir-next*: *SUCC*)
| *BytecodeExceptionNode* (*ir-arguments*: *INPUT list*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *ConditionalNode* (*ir-condition*: *INPUT-COND*) (*ir-trueValue*: *INPUT*) (*ir-falseValue*: *INPUT*)
| *ConstantNode* (*ir-const*: *Value*)
| *DynamicNewArrayNode* (*ir-elementType*: *INPUT*) (*ir-length*: *INPUT*) (*ir-voidClass-opt*: *INPUT option*) (*ir-stateBefore-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *EndNode*
| *ExceptionObjectNode* (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)

| *FrameState* (*ir-monitorIds*: *INPUT-ASSOC list*) (*ir-outerFrameState-opt*: *INPUT-STATE option*) (*ir-values-opt*: *INPUT list option*) (*ir-virtualObjectMappings-opt*: *INPUT-STATE list option*)
| *IfNode* (*ir-condition*: *INPUT-COND*) (*ir-trueSuccessor*: *SUCC*) (*ir-falseSuccessor*: *SUCC*)
| *IntegerBelowNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *IntegerEqualsNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)

| *IntegerLessThanNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *InvokeNode* (*ir-nid*: *ID*) (*ir-callTarget*: *INPUT-EXT*) (*ir-classInit-opt*: *INPUT option*) (*ir-stateDuring-opt*: *INPUT-STATE option*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *InvokeWithExceptionNode* (*ir-nid*: *ID*) (*ir-callTarget*: *INPUT-EXT*) (*ir-classInit-opt*: *INPUT option*) (*ir-stateDuring-opt*: *INPUT-STATE option*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*) (*ir-exceptionEdge*: *SUCC*)
| *IsNullNode* (*ir-value*: *INPUT*)
| *KillingBeginNode* (*ir-next*: *SUCC*)
| *LeftShiftNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *LoadFieldNode* (*ir-nid*: *ID*) (*ir-field*: *string*) (*ir-object-opt*: *INPUT option*) (*ir-next*: *SUCC*)
| *LogicNegationNode* (*ir-value*: *INPUT-COND*)
| *LoopBeginNode* (*ir-ends*: *INPUT-ASSOC list*) (*ir-overflowGuard-opt*: *INPUT-GUARD option*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *LoopEndNode* (*ir-loopBegin*: *INPUT-ASSOC*)
| *LoopExitNode* (*ir-loopBegin*: *INPUT-ASSOC*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *MergeNode* (*ir-ends*: *INPUT-ASSOC list*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *MethodCallTargetNode* (*ir-targetMethod*: *string*) (*ir-arguments*: *INPUT list*)
| *MulNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *NarrowNode* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*) (*ir-value*: *INPUT*)
| *NegateNode* (*ir-value*: *INPUT*)
| *NewArrayNode* (*ir-length*: *INPUT*) (*ir-stateBefore-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *NewInstanceNode* (*ir-nid*: *ID*) (*ir-instanceClass*: *string*) (*ir-stateBefore-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *NotNode* (*ir-value*: *INPUT*)
| *OrNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *ParameterNode* (*ir-index*: *nat*)
| *PiNode* (*ir-object*: *INPUT*) (*ir-guard-opt*: *INPUT-GUARD option*)
| *ReturnNode* (*ir-result-opt*: *INPUT option*) (*ir-memoryMap-opt*: *INPUT-EXT option*)
| *RightShiftNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *ShortCircuitOrNode* (*ir-x*: *INPUT-COND*) (*ir-y*: *INPUT-COND*)
| *SignExtendNode* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*) (*ir-value*: *INPUT*)
| *SignedDivNode* (*ir-nid*: *ID*) (*ir-x*: *INPUT*) (*ir-y*: *INPUT*) (*ir-zeroCheck-opt*: *INPUT-GUARD option*) (*ir-stateBefore-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)

| *SignedRemNode* (*ir-nid*: *ID*) (*ir-x*: *INPUT*) (*ir-y*: *INPUT*) (*ir-zeroCheck-opt*: *INPUT-GUARD option*) (*ir-stateBefore-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)

| *StartNode* (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-next*: *SUCC*)
| *StoreFieldNode* (*ir-nid*: *ID*) (*ir-field*: *string*) (*ir-value*: *INPUT*) (*ir-stateAfter-opt*: *INPUT-STATE option*) (*ir-object-opt*: *INPUT option*) (*ir-next*: *SUCC*)
| *SubNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *UnsignedRightShiftNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *UnwindNode* (*ir-exception*: *INPUT*)

| *ValuePhiNode* (*ir-nid*: *ID*) (*ir-values*: *INPUT list*) (*ir-merge*: *INPUT-ASSOC*)
| *ValueProxyNode* (*ir-value*: *INPUT*) (*ir-loopExit*: *INPUT-ASSOC*)
| *XorNode* (*ir-x*: *INPUT*) (*ir-y*: *INPUT*)
| *ZeroExtendNode* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*) (*ir-value*: *INPUT*)
| *NoNode*

| *RefNode* (*ir-ref*:*ID*)

fun *opt-to-list* :: *'a option* ⇒ *'a list* **where**
opt-to-list None = [] |
opt-to-list (*Some v*) = [*v*]

fun *opt-list-to-list* :: *'a list option* ⇒ *'a list* **where**
opt-list-to-list None = [] |
opt-list-to-list (*Some x*) = *x*

The following functions, inputs_of and successors_of, are automatically generated from the GraalVM compiler. Their purpose is to partition the node edges into input or successor edges.

fun *inputs-of* :: *IRNode* ⇒ *ID list* **where**
inputs-of-AbsNode:
inputs-of (*AbsNode value*) = [*value*] |
inputs-of-AddNode:
inputs-of (*AddNode x y*) = [*x*, *y*] |
inputs-of-AndNode:
inputs-of (*AndNode x y*) = [*x*, *y*] |
inputs-of-BeginNode:
inputs-of (*BeginNode next*) = [] |
inputs-of-BytecodeExceptionNode:
inputs-of (*BytecodeExceptionNode arguments stateAfter next*) = *arguments* @ (*opt-to-list stateAfter*) |
inputs-of-ConditionalNode:
inputs-of (*ConditionalNode condition trueValue falseValue*) = [*condition*, *trueValue*, *falseValue*] |
inputs-of-ConstantNode:
inputs-of (*ConstantNode const*) = [] |
inputs-of-DynamicNewArrayNode:
inputs-of (*DynamicNewArrayNode elementType length0 voidClass stateBefore next*) = [*elementType*, *length0*] @ (*opt-to-list voidClass*) @ (*opt-to-list stateBefore*) |
inputs-of-EndNode:
inputs-of (*EndNode*) = [] |
inputs-of-ExceptionObjectNode:
inputs-of (*ExceptionObjectNode stateAfter next*) = (*opt-to-list stateAfter*) |
inputs-of-FrameState:

inputs-of (*FrameState monitorIds outerFrameState values virtualObjectMappings*) = *monitorIds* @ (*opt-to-list outerFrameState*) @ (*opt-list-to-list values*) @ (*opt-list-to-list virtualObjectMappings*) |
inputs-of-IfNode:
inputs-of (*IfNode condition trueSuccessor falseSuccessor*) = [*condition*] |
inputs-of-IntegerBelowNode:
inputs-of (*IntegerBelowNode x y*) = [*x*, *y*] |
inputs-of-IntegerEqualsNode:
inputs-of (*IntegerEqualsNode x y*) = [*x*, *y*] |
inputs-of-IntegerLessThanNode:
inputs-of (*IntegerLessThanNode x y*) = [*x*, *y*] |
inputs-of-InvokeNode:
inputs-of (*InvokeNode nid0 callTarget classInit stateDuring stateAfter next*) = *callTarget* # (*opt-to-list classInit*) @ (*opt-to-list stateDuring*) @ (*opt-to-list stateAfter*) |
inputs-of-InvokeWithExceptionNode:
inputs-of (*InvokeWithExceptionNode nid0 callTarget classInit stateDuring stateAfter next exceptionEdge*) = *callTarget* # (*opt-to-list classInit*) @ (*opt-to-list stateDuring*) @ (*opt-to-list stateAfter*) |
inputs-of-IsNullNode:
inputs-of (*IsNullNode value*) = [*value*] |
inputs-of-KillingBeginNode:
inputs-of (*KillingBeginNode next*) = [] |
inputs-of-LeftShiftNode:
inputs-of (*LeftShiftNode x y*) = [*x*, *y*] |
inputs-of-LoadFieldNode:
inputs-of (*LoadFieldNode nid0 field object next*) = (*opt-to-list object*) |
inputs-of-LogicNegationNode:
inputs-of (*LogicNegationNode value*) = [*value*] |
inputs-of-LoopBeginNode:
inputs-of (*LoopBeginNode ends overflowGuard stateAfter next*) = *ends* @ (*opt-to-list overflowGuard*) @ (*opt-to-list stateAfter*) |
inputs-of-LoopEndNode:
inputs-of (*LoopEndNode loopBegin*) = [*loopBegin*] |
inputs-of-LoopExitNode:
inputs-of (*LoopExitNode loopBegin stateAfter next*) = *loopBegin* # (*opt-to-list stateAfter*) |
inputs-of-MergeNode:
inputs-of (*MergeNode ends stateAfter next*) = *ends* @ (*opt-to-list stateAfter*) |
inputs-of-MethodCallTargetNode:
inputs-of (*MethodCallTargetNode targetMethod arguments*) = *arguments* |
inputs-of-MulNode:
inputs-of (*MulNode x y*) = [*x*, *y*] |
inputs-of-NarrowNode:
inputs-of (*NarrowNode inputBits resultBits value*) = [*value*] |
inputs-of-NegateNode:
inputs-of (*NegateNode value*) = [*value*] |
inputs-of-NewArrayNode:
inputs-of (*NewArrayNode length0 stateBefore next*) = *length0* # (*opt-to-list state-*

Before) |
inputs-of-NewInstanceNode:
inputs-of (*NewInstanceNode nid0 instanceClass stateBefore next*) = (*opt-to-list stateBefore*) |
inputs-of-NotNode:
inputs-of (*NotNode value*) = [*value*] |
inputs-of-OrNode:
inputs-of (*OrNode x y*) = [*x*, *y*] |
inputs-of-ParameterNode:
inputs-of (*ParameterNode index*) = [] |
inputs-of-PiNode:
inputs-of (*PiNode object guard*) = *object* # (*opt-to-list guard*) |
inputs-of-ReturnNode:
inputs-of (*ReturnNode result memoryMap*) = (*opt-to-list result*) @ (*opt-to-list memoryMap*) |
inputs-of-RightShiftNode:
inputs-of (*RightShiftNode x y*) = [*x*, *y*] |
inputs-of-ShortCircuitOrNode:
inputs-of (*ShortCircuitOrNode x y*) = [*x*, *y*] |
inputs-of-SignExtendNode:
inputs-of (*SignExtendNode inputBits resultBits value*) = [*value*] |
inputs-of-SignedDivNode:
inputs-of (*SignedDivNode nid0 x y zeroCheck stateBefore next*) = [*x*, *y*] @ (*opt-to-list zeroCheck*) @ (*opt-to-list stateBefore*) |
inputs-of-SignedRemNode:
inputs-of (*SignedRemNode nid0 x y zeroCheck stateBefore next*) = [*x*, *y*] @ (*opt-to-list zeroCheck*) @ (*opt-to-list stateBefore*) |
inputs-of-StartNode:
inputs-of (*StartNode stateAfter next*) = (*opt-to-list stateAfter*) |
inputs-of-StoreFieldNode:
inputs-of (*StoreFieldNode nid0 field value stateAfter object next*) = *value* # (*opt-to-list stateAfter*) @ (*opt-to-list object*) |
inputs-of-SubNode:
inputs-of (*SubNode x y*) = [*x*, *y*] |
inputs-of-UnsignedRightShiftNode:
inputs-of (*UnsignedRightShiftNode x y*) = [*x*, *y*] |
inputs-of-UnwindNode:
inputs-of (*UnwindNode exception*) = [*exception*] |
inputs-of-ValuePhiNode:
inputs-of (*ValuePhiNode nid0 values merge*) = *merge* # *values* |
inputs-of-ValueProxyNode:
inputs-of (*ValueProxyNode value loopExit*) = [*value*, *loopExit*] |
inputs-of-XorNode:
inputs-of (*XorNode x y*) = [*x*, *y*] |
inputs-of-ZeroExtendNode:
inputs-of (*ZeroExtendNode inputBits resultBits value*) = [*value*] |
inputs-of-NoNode: *inputs-of* (*NoNode*) = [] |

inputs-of-RefNode: *inputs-of* (*RefNode ref*) = [*ref*]

fun *successors-of* :: *IRNode* ⇒ *ID list* **where**
successors-of-AbsNode:
successors-of (*AbsNode value*) = [] |
successors-of-AddNode:
successors-of (*AddNode x y*) = [] |
successors-of-AndNode:
successors-of (*AndNode x y*) = [] |
successors-of-BeginNode:
successors-of (*BeginNode next*) = [*next*] |
successors-of-BytecodeExceptionNode:
successors-of (*BytecodeExceptionNode arguments stateAfter next*) = [*next*] |
successors-of-ConditionalNode:
successors-of (*ConditionalNode condition trueValue falseValue*) = [] |
successors-of-ConstantNode:
successors-of (*ConstantNode const*) = [] |
successors-of-DynamicNewArrayNode:
successors-of (*DynamicNewArrayNode elementType length0 voidClass stateBefore next*) = [*next*] |
successors-of-EndNode:
successors-of (*EndNode*) = [] |
successors-of-ExceptionObjectNode:
successors-of (*ExceptionObjectNode stateAfter next*) = [*next*] |
successors-of-FrameState:
successors-of (*FrameState monitorIds outerFrameState values virtualObjectMappings*) = [] |
successors-of-IfNode:
successors-of (*IfNode condition trueSuccessor falseSuccessor*) = [*trueSuccessor*, *falseSuccessor*] |
successors-of-IntegerBelowNode:
successors-of (*IntegerBelowNode x y*) = [] |
successors-of-IntegerEqualsNode:
successors-of (*IntegerEqualsNode x y*) = [] |
successors-of-IntegerLessThanNode:
successors-of (*IntegerLessThanNode x y*) = [] |
successors-of-InvokeNode:
successors-of (*InvokeNode nid0 callTarget classInit stateDuring stateAfter next*) = [*next*] |
successors-of-InvokeWithExceptionNode:
successors-of (*InvokeWithExceptionNode nid0 callTarget classInit stateDuring stateAfter next exceptionEdge*) = [*next*, *exceptionEdge*] |
successors-of-IsNullNode:
successors-of (*IsNullNode value*) = [] |
successors-of-KillingBeginNode:
successors-of (*KillingBeginNode next*) = [*next*] |
successors-of-LeftShiftNode:
successors-of (*LeftShiftNode x y*) = [] |

$successors\text{-}of\text{-}LoadFieldNode$:
$successors\text{-}of\ (LoadFieldNode\ nid0\ field\ object\ next) = [next] \mid$
$successors\text{-}of\text{-}LogicNegationNode$:
$successors\text{-}of\ (LogicNegationNode\ value) = [] \mid$
$successors\text{-}of\text{-}LoopBeginNode$:
$successors\text{-}of\ (LoopBeginNode\ ends\ overflowGuard\ stateAfter\ next) = [next] \mid$
$successors\text{-}of\text{-}LoopEndNode$:
$successors\text{-}of\ (LoopEndNode\ loopBegin) = [] \mid$
$successors\text{-}of\text{-}LoopExitNode$:
$successors\text{-}of\ (LoopExitNode\ loopBegin\ stateAfter\ next) = [next] \mid$
$successors\text{-}of\text{-}MergeNode$:
$successors\text{-}of\ (MergeNode\ ends\ stateAfter\ next) = [next] \mid$
$successors\text{-}of\text{-}MethodCallTargetNode$:
$successors\text{-}of\ (MethodCallTargetNode\ targetMethod\ arguments) = [] \mid$
$successors\text{-}of\text{-}MulNode$:
$successors\text{-}of\ (MulNode\ x\ y) = [] \mid$
$successors\text{-}of\text{-}NarrowNode$:
$successors\text{-}of\ (NarrowNode\ inputBits\ resultBits\ value) = [] \mid$
$successors\text{-}of\text{-}NegateNode$:
$successors\text{-}of\ (NegateNode\ value) = [] \mid$
$successors\text{-}of\text{-}NewArrayNode$:
$successors\text{-}of\ (NewArrayNode\ length0\ stateBefore\ next) = [next] \mid$
$successors\text{-}of\text{-}NewInstanceNode$:
$successors\text{-}of\ (NewInstanceNode\ nid0\ instanceClass\ stateBefore\ next) = [next] \mid$
$successors\text{-}of\text{-}NotNode$:
$successors\text{-}of\ (NotNode\ value) = [] \mid$
$successors\text{-}of\text{-}OrNode$:
$successors\text{-}of\ (OrNode\ x\ y) = [] \mid$
$successors\text{-}of\text{-}ParameterNode$:
$successors\text{-}of\ (ParameterNode\ index) = [] \mid$
$successors\text{-}of\text{-}PiNode$:
$successors\text{-}of\ (PiNode\ object\ guard) = [] \mid$
$successors\text{-}of\text{-}ReturnNode$:
$successors\text{-}of\ (ReturnNode\ result\ memoryMap) = [] \mid$
$successors\text{-}of\text{-}RightShiftNode$:
$successors\text{-}of\ (RightShiftNode\ x\ y) = [] \mid$
$successors\text{-}of\text{-}ShortCircuitOrNode$:
$successors\text{-}of\ (ShortCircuitOrNode\ x\ y) = [] \mid$
$successors\text{-}of\text{-}SignExtendNode$:
$successors\text{-}of\ (SignExtendNode\ inputBits\ resultBits\ value) = [] \mid$
$successors\text{-}of\text{-}SignedDivNode$:
$successors\text{-}of\ (SignedDivNode\ nid0\ x\ y\ zeroCheck\ stateBefore\ next) = [next] \mid$
$successors\text{-}of\text{-}SignedRemNode$:
$successors\text{-}of\ (SignedRemNode\ nid0\ x\ y\ zeroCheck\ stateBefore\ next) = [next] \mid$
$successors\text{-}of\text{-}StartNode$:
$successors\text{-}of\ (StartNode\ stateAfter\ next) = [next] \mid$
$successors\text{-}of\text{-}StoreFieldNode$:
$successors\text{-}of\ (StoreFieldNode\ nid0\ field\ value\ stateAfter\ object\ next) = [next] \mid$
$successors\text{-}of\text{-}SubNode$:

successors-of (*SubNode x y*) = [] |
successors-of-UnsignedRightShiftNode:
successors-of (*UnsignedRightShiftNode x y*) = [] |
successors-of-UnwindNode:
successors-of (*UnwindNode exception*) = [] |
successors-of-ValuePhiNode:
successors-of (*ValuePhiNode nid0 values merge*) = [] |
successors-of-ValueProxyNode:
successors-of (*ValueProxyNode value loopExit*) = [] |
successors-of-XorNode:
successors-of (*XorNode x y*) = [] |
successors-of-ZeroExtendNode:
successors-of (*ZeroExtendNode inputBits resultBits value*) = [] |
successors-of-NoNode: *successors-of* (*NoNode*) = [] |

successors-of-RefNode: *successors-of* (*RefNode ref*) = [*ref*]

lemma *inputs-of* (*FrameState x* (*Some y*) (*Some z*) *None*) = *x* @ [*y*] @ *z*
⟨*proof*⟩
lemma *successors-of* (*FrameState x* (*Some y*) (*Some z*) *None*) = []
⟨*proof*⟩

lemma *inputs-of* (*IfNode c t f*) = [*c*]
⟨*proof*⟩
lemma *successors-of* (*IfNode c t f*) = [*t*, *f*]
⟨*proof*⟩

lemma *inputs-of* (*EndNode*) = [] ∧ *successors-of* (*EndNode*) = []
⟨*proof*⟩

end

3.2 IR Graph Node Hierarchy

theory *IRNodeHierarchy*
imports *IRNodes*
begin

It is helpful to introduce a node hierarchy into our formalization. Often the GraalVM compiler relies on explicit type checks to determine which operations to perform on a given node, we try to mimic the same functionality by using a suite of predicate functions over the IRNode class to determine inheritance.

As one would expect, the function is<ClassName>Type will be true if the node parameter is a subclass of the ClassName within the GraalVM compiler.

These functions have been automatically generated from the compiler.

fun *is-EndNode* :: *IRNode* ⇒ *bool* **where**
is-EndNode EndNode = *True* |
is-EndNode - = *False*

fun *is-VirtualState* :: *IRNode* ⇒ *bool* **where**
is-VirtualState n = ((*is-FrameState n*))

fun *is-BinaryArithmeticNode* :: *IRNode* ⇒ *bool* **where**
is-BinaryArithmeticNode n = ((*is-AddNode n*) ∨ (*is-AndNode n*) ∨ (*is-MulNode n*) ∨ (*is-OrNode n*) ∨ (*is-SubNode n*) ∨ (*is-XorNode n*))

fun *is-ShiftNode* :: *IRNode* ⇒ *bool* **where**
is-ShiftNode n = ((*is-LeftShiftNode n*) ∨ (*is-RightShiftNode n*) ∨ (*is-UnsignedRightShiftNode n*))

fun *is-BinaryNode* :: *IRNode* ⇒ *bool* **where**
is-BinaryNode n = ((*is-BinaryArithmeticNode n*) ∨ (*is-ShiftNode n*))

fun *is-AbstractLocalNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractLocalNode n = ((*is-ParameterNode n*))

fun *is-IntegerConvertNode* :: *IRNode* ⇒ *bool* **where**
is-IntegerConvertNode n = ((*is-NarrowNode n*) ∨ (*is-SignExtendNode n*) ∨ (*is-ZeroExtendNode n*))

fun *is-UnaryArithmeticNode* :: *IRNode* ⇒ *bool* **where**
is-UnaryArithmeticNode n = ((*is-AbsNode n*) ∨ (*is-NegateNode n*) ∨ (*is-NotNode n*))

fun *is-UnaryNode* :: *IRNode* ⇒ *bool* **where**
is-UnaryNode n = ((*is-IntegerConvertNode n*) ∨ (*is-UnaryArithmeticNode n*))

fun *is-PhiNode* :: *IRNode* ⇒ *bool* **where**
is-PhiNode n = ((*is-ValuePhiNode n*))

fun *is-FloatingGuardedNode* :: *IRNode* ⇒ *bool* **where**
is-FloatingGuardedNode n = ((*is-PiNode n*))

fun *is-UnaryOpLogicNode* :: *IRNode* ⇒ *bool* **where**
is-UnaryOpLogicNode n = ((*is-IsNullNode n*))

fun *is-IntegerLowerThanNode* :: *IRNode* ⇒ *bool* **where**
is-IntegerLowerThanNode n = ((*is-IntegerBelowNode n*) ∨ (*is-IntegerLessThanNode n*))

fun *is-CompareNode* :: *IRNode* ⇒ *bool* **where**
is-CompareNode n = ((*is-IntegerEqualsNode n*) ∨ (*is-IntegerLowerThanNode n*))

fun *is-BinaryOpLogicNode* :: *IRNode* ⇒ *bool* **where**
is-BinaryOpLogicNode n = ((*is-CompareNode n*))

fun *is-LogicNode* :: *IRNode* ⇒ *bool* **where**
is-LogicNode n = ((*is-BinaryOpLogicNode n*) ∨ (*is-LogicNegationNode n*) ∨ (*is-ShortCircuitOrNode n*) ∨ (*is-UnaryOpLogicNode n*))

fun *is-ProxyNode* :: *IRNode* ⇒ *bool* **where**
is-ProxyNode n = ((*is-ValueProxyNode n*))

fun *is-FloatingNode* :: *IRNode* ⇒ *bool* **where**
is-FloatingNode n = ((*is-AbstractLocalNode n*) ∨ (*is-BinaryNode n*) ∨ (*is-ConditionalNode n*) ∨ (*is-ConstantNode n*) ∨ (*is-FloatingGuardedNode n*) ∨ (*is-LogicNode n*) ∨ (*is-PhiNode n*) ∨ (*is-ProxyNode n*) ∨ (*is-UnaryNode n*))

fun *is-AccessFieldNode* :: *IRNode* ⇒ *bool* **where**
is-AccessFieldNode n = ((*is-LoadFieldNode n*) ∨ (*is-StoreFieldNode n*))

fun *is-AbstractNewArrayNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractNewArrayNode n = ((*is-DynamicNewArrayNode n*) ∨ (*is-NewArrayNode n*))

fun *is-AbstractNewObjectNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractNewObjectNode n = ((*is-AbstractNewArrayNode n*) ∨ (*is-NewInstanceNode n*))

fun *is-IntegerDivRemNode* :: *IRNode* ⇒ *bool* **where**
is-IntegerDivRemNode n = ((*is-SignedDivNode n*) ∨ (*is-SignedRemNode n*))

fun *is-FixedBinaryNode* :: *IRNode* ⇒ *bool* **where**
is-FixedBinaryNode n = ((*is-IntegerDivRemNode n*))

fun *is-DeoptimizingFixedWithNextNode* :: *IRNode* ⇒ *bool* **where**
is-DeoptimizingFixedWithNextNode n = ((*is-AbstractNewObjectNode n*) ∨ (*is-FixedBinaryNode n*))

fun *is-AbstractMemoryCheckpoint* :: *IRNode* ⇒ *bool* **where**
is-AbstractMemoryCheckpoint n = ((*is-BytecodeExceptionNode n*) ∨ (*is-InvokeNode n*))

fun *is-AbstractStateSplit* :: *IRNode* ⇒ *bool* **where**
is-AbstractStateSplit n = ((*is-AbstractMemoryCheckpoint n*))

fun *is-AbstractMergeNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractMergeNode n = ((*is-LoopBeginNode n*) ∨ (*is-MergeNode n*))

fun *is-BeginStateSplitNode* :: *IRNode* ⇒ *bool* **where**
is-BeginStateSplitNode n = ((*is-AbstractMergeNode n*) ∨ (*is-ExceptionObjectNode*

n) ∨ (*is-LoopExitNode* n) ∨ (*is-StartNode* n))

fun *is-AbstractBeginNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractBeginNode n = ((*is-BeginNode* n) ∨ (*is-BeginStateSplitNode* n) ∨ (*is-KillingBeginNode* n))

fun *is-FixedWithNextNode* :: *IRNode* ⇒ *bool* **where**
is-FixedWithNextNode n = ((*is-AbstractBeginNode* n) ∨ (*is-AbstractStateSplit* n) ∨ (*is-AccessFieldNode* n) ∨ (*is-DeoptimizingFixedWithNextNode* n))

fun *is-WithExceptionNode* :: *IRNode* ⇒ *bool* **where**
is-WithExceptionNode n = ((*is-InvokeWithExceptionNode* n))

fun *is-ControlSplitNode* :: *IRNode* ⇒ *bool* **where**
is-ControlSplitNode n = ((*is-IfNode* n) ∨ (*is-WithExceptionNode* n))

fun *is-ControlSinkNode* :: *IRNode* ⇒ *bool* **where**
is-ControlSinkNode n = ((*is-ReturnNode* n) ∨ (*is-UnwindNode* n))

fun *is-AbstractEndNode* :: *IRNode* ⇒ *bool* **where**
is-AbstractEndNode n = ((*is-EndNode* n) ∨ (*is-LoopEndNode* n))

fun *is-FixedNode* :: *IRNode* ⇒ *bool* **where**
is-FixedNode n = ((*is-AbstractEndNode* n) ∨ (*is-ControlSinkNode* n) ∨ (*is-ControlSplitNode* n) ∨ (*is-FixedWithNextNode* n))

fun *is-CallTargetNode* :: *IRNode* ⇒ *bool* **where**
is-CallTargetNode n = ((*is-MethodCallTargetNode* n))

fun *is-ValueNode* :: *IRNode* ⇒ *bool* **where**
is-ValueNode n = ((*is-CallTargetNode* n) ∨ (*is-FixedNode* n) ∨ (*is-FloatingNode* n))

fun *is-Node* :: *IRNode* ⇒ *bool* **where**
is-Node n = ((*is-ValueNode* n) ∨ (*is-VirtualState* n))

fun *is-MemoryKill* :: *IRNode* ⇒ *bool* **where**
is-MemoryKill n = ((*is-AbstractMemoryCheckpoint* n))

fun *is-NarrowableArithmeticNode* :: *IRNode* ⇒ *bool* **where**
is-NarrowableArithmeticNode n = ((*is-AbsNode* n) ∨ (*is-AddNode* n) ∨ (*is-AndNode* n) ∨ (*is-MulNode* n) ∨ (*is-NegateNode* n) ∨ (*is-NotNode* n) ∨ (*is-OrNode* n) ∨ (*is-ShiftNode* n) ∨ (*is-SubNode* n) ∨ (*is-XorNode* n))

fun *is-AnchoringNode* :: *IRNode* ⇒ *bool* **where**
is-AnchoringNode n = ((*is-AbstractBeginNode* n))

fun *is-DeoptBefore* :: *IRNode* ⇒ *bool* **where**
is-DeoptBefore n = ((*is-DeoptimizingFixedWithNextNode* n))

fun *is-IndirectCanonicalization* :: *IRNode* ⇒ *bool* **where**
is-IndirectCanonicalization n = ((*is-LogicNode n*))

fun *is-IterableNodeType* :: *IRNode* ⇒ *bool* **where**
is-IterableNodeType n = ((*is-AbstractBeginNode n*) ∨ (*is-AbstractMergeNode n*) ∨ (*is-FrameState n*) ∨ (*is-IfNode n*) ∨ (*is-IntegerDivRemNode n*) ∨ (*is-InvokeWithExceptionNode n*) ∨ (*is-LoopBeginNode n*) ∨ (*is-LoopExitNode n*) ∨ (*is-MethodCallTargetNode n*) ∨ (*is-ParameterNode n*) ∨ (*is-ReturnNode n*) ∨ (*is-ShortCircuitOrNode n*))

fun *is-Invoke* :: *IRNode* ⇒ *bool* **where**
is-Invoke n = ((*is-InvokeNode n*) ∨ (*is-InvokeWithExceptionNode n*))

fun *is-Proxy* :: *IRNode* ⇒ *bool* **where**
is-Proxy n = ((*is-ProxyNode n*))

fun *is-ValueProxy* :: *IRNode* ⇒ *bool* **where**
is-ValueProxy n = ((*is-PiNode n*) ∨ (*is-ValueProxyNode n*))

fun *is-ValueNodeInterface* :: *IRNode* ⇒ *bool* **where**
is-ValueNodeInterface n = ((*is-ValueNode n*))

fun *is-ArrayLengthProvider* :: *IRNode* ⇒ *bool* **where**
is-ArrayLengthProvider n = ((*is-AbstractNewArrayNode n*) ∨ (*is-ConstantNode n*))

fun *is-StampInverter* :: *IRNode* ⇒ *bool* **where**
is-StampInverter n = ((*is-IntegerConvertNode n*) ∨ (*is-NegateNode n*) ∨ (*is-NotNode n*))

fun *is-GuardingNode* :: *IRNode* ⇒ *bool* **where**
is-GuardingNode n = ((*is-AbstractBeginNode n*))

fun *is-SingleMemoryKill* :: *IRNode* ⇒ *bool* **where**
is-SingleMemoryKill n = ((*is-BytecodeExceptionNode n*) ∨ (*is-ExceptionObjectNode n*) ∨ (*is-InvokeNode n*) ∨ (*is-InvokeWithExceptionNode n*) ∨ (*is-KillingBeginNode n*) ∨ (*is-StartNode n*))

fun *is-LIRLowerable* :: *IRNode* ⇒ *bool* **where**
is-LIRLowerable n = ((*is-AbstractBeginNode n*) ∨ (*is-AbstractEndNode n*) ∨ (*is-AbstractMergeNode n*) ∨ (*is-BinaryOpLogicNode n*) ∨ (*is-CallTargetNode n*) ∨ (*is-ConditionalNode n*) ∨ (*is-ConstantNode n*) ∨ (*is-IfNode n*) ∨ (*is-InvokeNode n*) ∨ (*is-InvokeWithExceptionNode n*) ∨ (*is-IsNullNode n*) ∨ (*is-LoopBeginNode n*) ∨ (*is-PiNode n*) ∨ (*is-ReturnNode n*) ∨ (*is-SignedDivNode n*) ∨ (*is-SignedRemNode n*) ∨ (*is-UnaryOpLogicNode n*) ∨ (*is-UnwindNode n*))

fun *is-GuardedNode* :: *IRNode* ⇒ *bool* **where**
is-GuardedNode n = ((*is-FloatingGuardedNode n*))

fun *is-ArithmeticLIRLowerable* :: *IRNode* ⇒ *bool* **where**
is-ArithmeticLIRLowerable n = ((*is-AbsNode n*) ∨ (*is-BinaryArithmeticNode n*) ∨ (*is-IntegerConvertNode n*) ∨ (*is-NotNode n*) ∨ (*is-ShiftNode n*) ∨ (*is-UnaryArithmeticNode n*))

fun *is-SwitchFoldable* :: *IRNode* ⇒ *bool* **where**
is-SwitchFoldable n = ((*is-IfNode n*))

fun *is-VirtualizableAllocation* :: *IRNode* ⇒ *bool* **where**
is-VirtualizableAllocation n = ((*is-NewArrayNode n*) ∨ (*is-NewInstanceNode n*))

fun *is-Unary* :: *IRNode* ⇒ *bool* **where**
is-Unary n = ((*is-LoadFieldNode n*) ∨ (*is-LogicNegationNode n*) ∨ (*is-UnaryNode n*) ∨ (*is-UnaryOpLogicNode n*))

fun *is-FixedNodeInterface* :: *IRNode* ⇒ *bool* **where**
is-FixedNodeInterface n = ((*is-FixedNode n*))

fun *is-BinaryCommutative* :: *IRNode* ⇒ *bool* **where**
is-BinaryCommutative n = ((*is-AddNode n*) ∨ (*is-AndNode n*) ∨ (*is-IntegerEqualsNode n*) ∨ (*is-MulNode n*) ∨ (*is-OrNode n*) ∨ (*is-XorNode n*))

fun *is-Canonicalizable* :: *IRNode* ⇒ *bool* **where**
is-Canonicalizable n = ((*is-BytecodeExceptionNode n*) ∨ (*is-ConditionalNode n*) ∨ (*is-DynamicNewArrayNode n*) ∨ (*is-PhiNode n*) ∨ (*is-PiNode n*) ∨ (*is-ProxyNode n*) ∨ (*is-StoreFieldNode n*) ∨ (*is-ValueProxyNode n*))

fun *is-UncheckedInterfaceProvider* :: *IRNode* ⇒ *bool* **where**
is-UncheckedInterfaceProvider n = ((*is-InvokeNode n*) ∨ (*is-InvokeWithExceptionNode n*) ∨ (*is-LoadFieldNode n*) ∨ (*is-ParameterNode n*))

fun *is-Binary* :: *IRNode* ⇒ *bool* **where**
is-Binary n = ((*is-BinaryArithmeticNode n*) ∨ (*is-BinaryNode n*) ∨ (*is-BinaryOpLogicNode n*) ∨ (*is-CompareNode n*) ∨ (*is-FixedBinaryNode n*) ∨ (*is-ShortCircuitOrNode n*))

fun *is-ArithmeticOperation* :: *IRNode* ⇒ *bool* **where**
is-ArithmeticOperation n = ((*is-BinaryArithmeticNode n*) ∨ (*is-IntegerConvertNode n*) ∨ (*is-ShiftNode n*) ∨ (*is-UnaryArithmeticNode n*))

fun *is-ValueNumberable* :: *IRNode* ⇒ *bool* **where**
is-ValueNumberable n = ((*is-FloatingNode n*) ∨ (*is-ProxyNode n*))

fun *is-Lowerable* :: *IRNode* ⇒ *bool* **where**
is-Lowerable n = ((*is-AbstractNewObjectNode n*) ∨ (*is-AccessFieldNode n*) ∨ (*is-BytecodeExceptionNode n*) ∨ (*is-ExceptionObjectNode n*) ∨ (*is-IntegerDivRemNode n*) ∨ (*is-UnwindNode n*))

fun *is-Virtualizable* :: *IRNode* ⇒ *bool* **where**
is-Virtualizable n = ((*is-IsNullNode n*) ∨ (*is-LoadFieldNode n*) ∨ (*is-PiNode n*)

$\vee$ (*is-StoreFieldNode n*) $\vee$ (*is-ValueProxyNode n*))

fun *is-Simplifiable* :: *IRNode* $\Rightarrow$ *bool* **where**
is-Simplifiable n = ((*is-AbstractMergeNode n*) $\vee$ (*is-BeginNode n*) $\vee$ (*is-IfNode n*) $\vee$ (*is-LoopExitNode n*) $\vee$ (*is-MethodCallTargetNode n*) $\vee$ (*is-NewArrayNode n*))

fun *is-StateSplit* :: *IRNode* $\Rightarrow$ *bool* **where**
is-StateSplit n = ((*is-AbstractStateSplit n*) $\vee$ (*is-BeginStateSplitNode n*) $\vee$ (*is-StoreFieldNode n*))

fun *is-ConvertNode* :: *IRNode* $\Rightarrow$ *bool* **where**
is-ConvertNode n = ((*is-IntegerConvertNode n*))

fun *is-sequential-node* :: *IRNode* $\Rightarrow$ *bool* **where**
is-sequential-node (*StartNode - -*) = *True* |
is-sequential-node (*BeginNode -*) = *True* |
is-sequential-node (*KillingBeginNode -*) = *True* |
is-sequential-node (*LoopBeginNode - - - -*) = *True* |
is-sequential-node (*LoopExitNode - - -*) = *True* |
is-sequential-node (*MergeNode - - -*) = *True* |
is-sequential-node (*RefNode -*) = *True* |
is-sequential-node - = *False*

The following convenience function is useful in determining if two IRNodes are of the same type irregardless of their edges. It will return true if both the node parameters are the same node class.

fun *is-same-ir-node-type* :: *IRNode* $\Rightarrow$ *IRNode* $\Rightarrow$ *bool* **where**
is-same-ir-node-type n1 n2 = (
((*is-AbsNode n1*) $\wedge$ (*is-AbsNode n2*)) $\vee$
((*is-AddNode n1*) $\wedge$ (*is-AddNode n2*)) $\vee$
((*is-AndNode n1*) $\wedge$ (*is-AndNode n2*)) $\vee$
((*is-BeginNode n1*) $\wedge$ (*is-BeginNode n2*)) $\vee$
((*is-BytecodeExceptionNode n1*) $\wedge$ (*is-BytecodeExceptionNode n2*)) $\vee$
((*is-ConditionalNode n1*) $\wedge$ (*is-ConditionalNode n2*)) $\vee$
((*is-ConstantNode n1*) $\wedge$ (*is-ConstantNode n2*)) $\vee$
((*is-DynamicNewArrayNode n1*) $\wedge$ (*is-DynamicNewArrayNode n2*)) $\vee$
((*is-EndNode n1*) $\wedge$ (*is-EndNode n2*)) $\vee$
((*is-ExceptionObjectNode n1*) $\wedge$ (*is-ExceptionObjectNode n2*)) $\vee$
((*is-FrameState n1*) $\wedge$ (*is-FrameState n2*)) $\vee$
((*is-IfNode n1*) $\wedge$ (*is-IfNode n2*)) $\vee$
((*is-IntegerBelowNode n1*) $\wedge$ (*is-IntegerBelowNode n2*)) $\vee$
((*is-IntegerEqualsNode n1*) $\wedge$ (*is-IntegerEqualsNode n2*)) $\vee$
((*is-IntegerLessThanNode n1*) $\wedge$ (*is-IntegerLessThanNode n2*)) $\vee$
((*is-InvokeNode n1*) $\wedge$ (*is-InvokeNode n2*)) $\vee$
((*is-InvokeWithExceptionNode n1*) $\wedge$ (*is-InvokeWithExceptionNode n2*)) $\vee$
((*is-IsNullNode n1*) $\wedge$ (*is-IsNullNode n2*)) $\vee$
((*is-KillingBeginNode n1*) $\wedge$ (*is-KillingBeginNode n2*)) $\vee$
((*is-LeftShiftNode n1*) $\wedge$ (*is-LeftShiftNode n2*)) $\vee$

$((\textit{is-LoadFieldNode n1}) \wedge (\textit{is-LoadFieldNode n2})) \vee$
$((\textit{is-LogicNegationNode n1}) \wedge (\textit{is-LogicNegationNode n2})) \vee$
$((\textit{is-LoopBeginNode n1}) \wedge (\textit{is-LoopBeginNode n2})) \vee$
$((\textit{is-LoopEndNode n1}) \wedge (\textit{is-LoopEndNode n2})) \vee$
$((\textit{is-LoopExitNode n1}) \wedge (\textit{is-LoopExitNode n2})) \vee$
$((\textit{is-MergeNode n1}) \wedge (\textit{is-MergeNode n2})) \vee$
$((\textit{is-MethodCallTargetNode n1}) \wedge (\textit{is-MethodCallTargetNode n2})) \vee$
$((\textit{is-MulNode n1}) \wedge (\textit{is-MulNode n2})) \vee$
$((\textit{is-NarrowNode n1}) \wedge (\textit{is-NarrowNode n2})) \vee$
$((\textit{is-NegateNode n1}) \wedge (\textit{is-NegateNode n2})) \vee$
$((\textit{is-NewArrayNode n1}) \wedge (\textit{is-NewArrayNode n2})) \vee$
$((\textit{is-NewInstanceNode n1}) \wedge (\textit{is-NewInstanceNode n2})) \vee$
$((\textit{is-NotNode n1}) \wedge (\textit{is-NotNode n2})) \vee$
$((\textit{is-OrNode n1}) \wedge (\textit{is-OrNode n2})) \vee$
$((\textit{is-ParameterNode n1}) \wedge (\textit{is-ParameterNode n2})) \vee$
$((\textit{is-PiNode n1}) \wedge (\textit{is-PiNode n2})) \vee$
$((\textit{is-ReturnNode n1}) \wedge (\textit{is-ReturnNode n2})) \vee$
$((\textit{is-RightShiftNode n1}) \wedge (\textit{is-RightShiftNode n2})) \vee$
$((\textit{is-ShortCircuitOrNode n1}) \wedge (\textit{is-ShortCircuitOrNode n2})) \vee$
$((\textit{is-SignedDivNode n1}) \wedge (\textit{is-SignedDivNode n2})) \vee$
$((\textit{is-SignedRemNode n1}) \wedge (\textit{is-SignedRemNode n2})) \vee$
$((\textit{is-SignExtendNode n1}) \wedge (\textit{is-SignExtendNode n2})) \vee$
$((\textit{is-StartNode n1}) \wedge (\textit{is-StartNode n2})) \vee$
$((\textit{is-StoreFieldNode n1}) \wedge (\textit{is-StoreFieldNode n2})) \vee$
$((\textit{is-SubNode n1}) \wedge (\textit{is-SubNode n2})) \vee$
$((\textit{is-UnsignedRightShiftNode n1}) \wedge (\textit{is-UnsignedRightShiftNode n2})) \vee$
$((\textit{is-UnwindNode n1}) \wedge (\textit{is-UnwindNode n2})) \vee$
$((\textit{is-ValuePhiNode n1}) \wedge (\textit{is-ValuePhiNode n2})) \vee$
$((\textit{is-ValueProxyNode n1}) \wedge (\textit{is-ValueProxyNode n2})) \vee$
$((\textit{is-XorNode n1}) \wedge (\textit{is-XorNode n2})) \vee$
$((\textit{is-ZeroExtendNode n1}) \wedge (\textit{is-ZeroExtendNode n2})))$

end

3.3 IR Graph Type

theory *IRGraph*
  imports
    IRNodeHierarchy
    Stamp
    HOL−Library.FSet
    HOL.Relation
begin

This theory defines the main Graal data structure - an entire IR Graph.

IRGraph is defined as a partial map with a finite domain. The finite domain is required to be able to generate code and produce an interpreter.

typedef $\textit{IRGraph} = \{g :: \textit{ID} \rightharpoonup (\textit{IRNode} \times \textit{Stamp}) \ . \ \textit{finite} \ (\textit{dom} \ g)\}$

$\langle proof \rangle$

setup-lifting *type-definition-IRGraph*

lift-definition *ids* :: *IRGraph* ⇒ *ID set*
  is $\lambda g.\ \{nid \in dom\ g\ .\ \nexists s.\ g\ nid = (Some\ (NoNode,\ s))\}$ $\langle proof \rangle$

fun *with-default* :: $'c \Rightarrow ('b \Rightarrow 'c) \Rightarrow (('a \rightharpoonup 'b) \Rightarrow 'a \Rightarrow 'c)$ **where**
  $with\text{-}default\ def\ conv = (\lambda m\ k.$
    $(case\ m\ k\ of\ None \Rightarrow def \mid Some\ v \Rightarrow conv\ v))$

lift-definition *kind* :: $IRGraph \Rightarrow (ID \Rightarrow IRNode)$
  is *with-default NoNode fst* $\langle proof \rangle$

lift-definition *stamp* :: $IRGraph \Rightarrow ID \Rightarrow Stamp$
  is *with-default IllegalStamp snd* $\langle proof \rangle$

lift-definition *add-node* :: $ID \Rightarrow (IRNode \times Stamp) \Rightarrow IRGraph \Rightarrow IRGraph$
  is $\lambda nid\ k\ g.\ if\ fst\ k = NoNode\ then\ g\ else\ g(nid \mapsto k)$ $\langle proof \rangle$

lift-definition *remove-node* :: $ID \Rightarrow IRGraph \Rightarrow IRGraph$
  is $\lambda nid\ g.\ g(nid := None)$ $\langle proof \rangle$

lift-definition *replace-node* :: $ID \Rightarrow (IRNode \times Stamp) \Rightarrow IRGraph \Rightarrow IRGraph$
  is $\lambda nid\ k\ g.\ if\ fst\ k = NoNode\ then\ g\ else\ g(nid \mapsto k)$ $\langle proof \rangle$

lift-definition *as-list* :: $IRGraph \Rightarrow (ID \times IRNode \times Stamp)\ list$
  is $\lambda g.\ map\ (\lambda k.\ (k,\ the\ (g\ k)))\ (sorted\text{-}list\text{-}of\text{-}set\ (dom\ g))$ $\langle proof \rangle$

fun *no-node* :: $(ID \times (IRNode \times Stamp))\ list \Rightarrow (ID \times (IRNode \times Stamp))\ list$
where
  $no\text{-}node\ g = filter\ (\lambda n.\ fst\ (snd\ n) \neq NoNode)\ g$

lift-definition *irgraph* :: $(ID \times (IRNode \times Stamp))\ list \Rightarrow IRGraph$
  is $map\text{-}of \circ no\text{-}node$
  $\langle proof \rangle$

definition *as-set* :: $IRGraph \Rightarrow (ID \times (IRNode \times Stamp))\ set$ **where**
  $as\text{-}set\ g = \{(n,\ kind\ g\ n,\ stamp\ g\ n) \mid n\ .\ n \in ids\ g\}$

definition *true-ids* :: $IRGraph \Rightarrow ID\ set$ **where**
  $true\text{-}ids\ g = ids\ g - \{n \in ids\ g.\ \exists\, n'\ .\ kind\ g\ n = RefNode\ n'\}$

definition *domain-subtraction* :: $'a\ set \Rightarrow ('a \times 'b)\ set \Rightarrow ('a \times 'b)\ set$
  (**infix** $\trianglelefteq$ *30*) **where**
  $domain\text{-}subtraction\ s\ r = \{(x,\ y)\ .\ (x,\ y) \in r \wedge x \notin s\}$

notation (*latex*)
  domain-subtraction (- $\lhd$ -)

code-datatype *irgraph*

fun *filter-none* **where**
  $filter\text{-}none\ g = \{nid \in dom\ g\ .\ \nexists s.\ g\ nid = (Some\ (NoNode,\ s))\}$

lemma *no-node-clears*:
  $res = no\text{-}node\ xs \longrightarrow (\forall x \in set\ res.\ fst\ (snd\ x) \neq NoNode)$
  ⟨*proof*⟩

lemma *dom-eq*:
  assumes $\forall x \in set\ xs.\ fst\ (snd\ x) \neq NoNode$
  shows $filter\text{-}none\ (map\text{-}of\ xs) = dom\ (map\text{-}of\ xs)$
  ⟨*proof*⟩

lemma *fil-eq*:
  $filter\text{-}none\ (map\text{-}of\ (no\text{-}node\ xs)) = set\ (map\ fst\ (no\text{-}node\ xs))$
  ⟨*proof*⟩

lemma *irgraph*[*code*]: $ids\ (irgraph\ m) = set\ (map\ fst\ (no\text{-}node\ m))$
  ⟨*proof*⟩

lemma [*code*]: $Rep\text{-}IRGraph\ (irgraph\ m) = map\text{-}of\ (no\text{-}node\ m)$
  ⟨*proof*⟩
fun *inputs* :: $IRGraph \Rightarrow ID \Rightarrow ID\ set$ **where**
  $inputs\ g\ nid = set\ (inputs\text{-}of\ (kind\ g\ nid))$
— Get the successor set of a given node ID
fun *succ* :: $IRGraph \Rightarrow ID \Rightarrow ID\ set$ **where**
  $succ\ g\ nid = set\ (successors\text{-}of\ (kind\ g\ nid))$
— Gives a relation between node IDs - between a node and its input nodes
fun *input-edges* :: $IRGraph \Rightarrow ID\ rel$ **where**
  $input\text{-}edges\ g = (\bigcup\ i \in ids\ g.\ \{(i,j)|j.\ j \in (inputs\ g\ i)\})$
— Find all the nodes in the graph that have nid as an input - the usages of nid
fun *usages* :: $IRGraph \Rightarrow ID \Rightarrow ID\ set$ **where**
  $usages\ g\ nid = \{i.\ i \in ids\ g \wedge nid \in inputs\ g\ i\}$
fun *successor-edges* :: $IRGraph \Rightarrow ID\ rel$ **where**
  $successor\text{-}edges\ g = (\bigcup\ i \in ids\ g.\ \{(i,j)|j\ .\ j \in (succ\ \ g\ i)\})$
fun *predecessors* :: $IRGraph \Rightarrow ID \Rightarrow ID\ set$ **where**
  $predecessors\ g\ nid = \{i.\ i \in ids\ g \wedge nid \in succ\ g\ i\}$
fun *nodes-of* :: $IRGraph \Rightarrow (IRNode \Rightarrow bool) \Rightarrow ID\ set$ **where**
  $nodes\text{-}of\ g\ sel = \{nid \in ids\ g\ .\ sel\ (kind\ g\ nid)\}$
fun *edge* :: $(IRNode \Rightarrow 'a) \Rightarrow ID \Rightarrow IRGraph \Rightarrow 'a$ **where**
  $edge\ sel\ nid\ g = sel\ (kind\ g\ nid)$

fun *filtered-inputs* :: $IRGraph \Rightarrow ID \Rightarrow (IRNode \Rightarrow bool) \Rightarrow ID\ list$ **where**
  $filtered\text{-}inputs\ g\ nid\ f = filter\ (f \circ (kind\ g))\ (inputs\text{-}of\ (kind\ g\ nid))$
fun *filtered-successors* :: $IRGraph \Rightarrow ID \Rightarrow (IRNode \Rightarrow bool) \Rightarrow ID\ list$ **where**
  $filtered\text{-}successors\ g\ nid\ f = filter\ (f \circ (kind\ g))\ (successors\text{-}of\ (kind\ g\ nid))$

fun *filtered-usages* :: *IRGraph* ⇒ *ID* ⇒ (*IRNode* ⇒ *bool*) ⇒ *ID set* **where**
filtered-usages g nid f = {*n* ∈ (*usages g nid*). *f* (*kind g n*)}

fun *is-empty* :: *IRGraph* ⇒ *bool* **where**
is-empty g = (*ids g* = {})

fun *any-usage* :: *IRGraph* ⇒ *ID* ⇒ *ID* **where**
any-usage g nid = *hd* (*sorted-list-of-set* (*usages g nid*))

lemma *ids-some*[*simp*]: *x* ∈ *ids g* ⟷ *kind g x* ≠ *NoNode*
⟨*proof*⟩

lemma *not-in-g*:
assumes *nid* ∉ *ids g*
shows *kind g nid* = *NoNode*
⟨*proof*⟩

lemma *valid-creation*[*simp*]:
finite (*dom g*) ⟷ *Rep-IRGraph* (*Abs-IRGraph g*) = *g*
⟨*proof*⟩

lemma [*simp*]: *finite* (*ids g*)
⟨*proof*⟩

lemma [*simp*]: *finite* (*ids* (*irgraph g*))
⟨*proof*⟩

lemma [*simp*]: *finite* (*dom g*) ⟶ *ids* (*Abs-IRGraph g*) = {*nid* ∈ *dom g* . ∄ *s*. *g nid* = *Some* (*NoNode*, *s*)}
⟨*proof*⟩

lemma [*simp*]: *finite* (*dom g*) ⟶ *kind* (*Abs-IRGraph g*) = (λ*x* . (*case g x of None* ⇒ *NoNode* | *Some n* ⇒ *fst n*))
⟨*proof*⟩

lemma [*simp*]: *finite* (*dom g*) ⟶ *stamp* (*Abs-IRGraph g*) = (λ*x* . (*case g x of None* ⇒ *IllegalStamp* | *Some n* ⇒ *snd n*))
⟨*proof*⟩

lemma [*simp*]: *ids* (*irgraph g*) = *set* (*map fst* (*no-node g*))
⟨*proof*⟩

lemma [*simp*]: *kind* (*irgraph g*) = (λ*nid*. (*case* (*map-of* (*no-node g*)) *nid of None* ⇒ *NoNode* | *Some n* ⇒ *fst n*))
⟨*proof*⟩

lemma [*simp*]: *stamp* (*irgraph g*) = (λ*nid*. (*case* (*map-of* (*no-node g*)) *nid of None* ⇒ *IllegalStamp* | *Some n* ⇒ *snd n*))
⟨*proof*⟩

lemma *map-of-upd*: $(\textit{map-of } g)(k \mapsto v) = (\textit{map-of } ((k, v) \# g))$
⟨*proof*⟩

lemma [*code*]: $\textit{replace-node nid } k\ (\textit{irgraph } g) = (\textit{irgraph } (\ ((\textit{nid}, k) \# g)))$
⟨*proof*⟩

lemma [*code*]: $\textit{add-node nid } k\ (\textit{irgraph } g) = (\textit{irgraph } (((\textit{nid}, k) \# g)))$
⟨*proof*⟩

lemma *add-node-lookup*:
$\textit{gup} = \textit{add-node nid } (k, s)\ g \longrightarrow$
$(\textit{if } k \neq \textit{NoNode then kind gup nid} = k \wedge \textit{stamp gup nid} = s \textit{ else kind gup nid} = \textit{kind } g \textit{ nid})$
⟨*proof*⟩

lemma *remove-node-lookup*:
$\textit{gup} = \textit{remove-node nid } g \longrightarrow \textit{kind gup nid} = \textit{NoNode} \wedge \textit{stamp gup nid} = \textit{IllegalStamp}$
⟨*proof*⟩

lemma *replace-node-lookup*[*simp*]:
$\textit{gup} = \textit{replace-node nid } (k, s)\ g \wedge k \neq \textit{NoNode} \longrightarrow \textit{kind gup nid} = k \wedge \textit{stamp gup nid} = s$
⟨*proof*⟩

lemma *replace-node-unchanged*:
$\textit{gup} = \textit{replace-node nid } (k, s)\ g \longrightarrow (\forall\ n \in (\textit{ids } g - \{\textit{nid}\})\ .\ n \in \textit{ids } g \wedge n \in \textit{ids gup} \wedge \textit{kind } g\ n = \textit{kind gup } n)$
⟨*proof*⟩

3.3.1 Example Graphs

Example 1: empty graph (just a start and end node)

definition *start-end-graph*:: *IRGraph* **where**
start-end-graph = *irgraph* [(*0*, *StartNode None 1*, *VoidStamp*), (*1*, *ReturnNode None None*, *VoidStamp*)]

Example 2: public static int sq(int x) return x * x;

[1 P(0)] / [0 Start] [4 *] | / V / [5 Return]

definition *eg2-sq* :: *IRGraph* **where**
eg2-sq = *irgraph* [
(*0*, *StartNode None 5*, *VoidStamp*),
(*1*, *ParameterNode 0*, *default-stamp*),
(*4*, *MulNode 1 1*, *default-stamp*),
(*5*, *ReturnNode* (*Some 4*) *None*, *default-stamp*)
]

value *input-edges eg2-sq*
value *usages eg2-sq 1*

end

3.4 Control-flow Graph Traversal

theory
Traversal
imports
IRGraph
begin

type-synonym *Seen* = *ID set*

nextEdge helps determine which node to traverse next by returning the first successor edge that isn't in the set of already visited nodes. If there is not an appropriate successor, None is returned instead.

fun *nextEdge* :: *Seen* ⇒ *ID* ⇒ *IRGraph* ⇒ *ID option* **where**
nextEdge seen nid g =
(*let nids* = (*filter* (λ*nid*′. *nid*′ ∉ *seen*) (*successors-of* (*kind g nid*))) *in*
(*if length nids* > *0 then Some* (*hd nids*) *else None*))

pred determines which node, if any, acts as the predecessor of another.

Merge nodes represent a special case where-in the predecessor exists as an input edge of the merge node, to simplify the traversal we treat only the first input end node as the predecessor, ignoring that multiple nodes may act as a successor.

For all other nodes, the predecessor is the first element of the predecessors set. Note that in a well-formed graph there should only be one element in the predecessor set.

fun *pred* :: *IRGraph* ⇒ *ID* ⇒ *ID option* **where**
pred g nid = (*case kind g nid of*
(*MergeNode ends - -*) ⇒ *Some* (*hd ends*) |
\- ⇒
(*if IRGraph.predecessors g nid* = {}
then None else
Some (*hd* (*sorted-list-of-set* (*IRGraph.predecessors g nid*)))
)
)

Here we try to implement a generic fork of the control-flow traversal algorithm that was initially implemented for the ConditionalElimination phase

type-synonym $'a$ *TraversalState* $= (ID \times Seen \times 'a)$

inductive *Step*
$:: ('a\ TraversalState \Rightarrow 'a) \Rightarrow IRGraph \Rightarrow 'a\ TraversalState \Rightarrow 'a\ TraversalState\ option \Rightarrow bool$
for *sa g* **where**
— Hit a BeginNode with an IfNode predecessor which represents the start of a basic block for the IfNode. 1. nid' will be the successor of the begin node. 2. Find the first and only predecessor. 3. Extract condition from the preceding IfNode. 4. Negate condition if the begin node is second branch (we've taken the else branch of the condition) 5. Add the condition or the negated condition to stack 6. Perform any stamp updates based on the condition using the registerNewCondition function and place them on the top of the stack of stamp information
$[\![kind\ g\ nid = BeginNode\ nid';$

$nid \notin seen;$
$seen' = \{nid\} \cup seen;$

$Some\ ifcond = pred\ g\ nid;$
$kind\ g\ ifcond = IfNode\ cond\ t\ f;$

$analysis' = sa\ (nid, seen, analysis)]\!]$
$\Longrightarrow Step\ sa\ g\ (nid, seen, analysis)\ (Some\ (nid', seen', analysis'))\ |$

— Hit an EndNode 1. nid' will be the usage of EndNode 2. pop the conditions and stamp stack
$[\![kind\ g\ nid = EndNode;$

$nid \notin seen;$
$seen' = \{nid\} \cup seen;$

$nid' = any\text{-}usage\ g\ nid;$

$analysis' = sa\ (nid, seen, analysis)]\!]$
$\Longrightarrow Step\ sa\ g\ (nid, seen, analysis)\ (Some\ (nid', seen', analysis'))\ |$

— We can find a successor edge that is not in seen, go there
$[\![\neg(is\text{-}EndNode\ (kind\ g\ nid));$
$\neg(is\text{-}BeginNode\ (kind\ g\ nid));$

$nid \notin seen;$
$seen' = \{nid\} \cup seen;$

$Some\ nid' = nextEdge\ seen'\ nid\ g;$

$analysis' = sa\ (nid, seen, analysis)]\!]$
$\Longrightarrow Step\ sa\ g\ (nid, seen, analysis)\ (Some\ (nid', seen', analysis'))\ |$

— We can cannot find a successor edge that is not in seen, give back None

$\llbracket\neg(\textit{is-EndNode}\ (\textit{kind}\ g\ \textit{nid}));$
$\neg(\textit{is-BeginNode}\ (\textit{kind}\ g\ \textit{nid}));$

$\textit{nid} \notin \textit{seen};$
$\textit{seen}' = \{\textit{nid}\} \cup \textit{seen};$

$\textit{None} = \textit{nextEdge}\ \textit{seen}'\ \textit{nid}\ g\rrbracket$
$\Longrightarrow \textit{Step}\ \textit{sa}\ g\ (\textit{nid}, \textit{seen}, \textit{analysis})\ \textit{None}\ |$

— We've already seen this node, give back None
$\llbracket \textit{nid} \in \textit{seen}\rrbracket \Longrightarrow \textit{Step}\ \textit{sa}\ g\ (\textit{nid}, \textit{seen}, \textit{analysis})\ \textit{None}$

code-pred $(\textit{modes}\colon i \Rightarrow i \Rightarrow i \Rightarrow o \Rightarrow \textit{bool})\ \textit{Step}\ \langle \textit{proof}\rangle$

end

3.5 Structural Graph Comparison

theory
Comparison
imports
IRGraph
begin

We introduce a form of structural graph comparison that is able to assert structural equivalence of graphs which differ in zero or more reference node chains for any given nodes.

fun $\textit{find-ref-nodes} :: \textit{IRGraph} \Rightarrow (\textit{ID} \rightharpoonup \textit{ID})$ **where**
$\textit{find-ref-nodes}\ g = \textit{map-of}$
$(\textit{map}\ (\lambda n.\ (n, \textit{ir-ref}\ (\textit{kind}\ g\ n)))\ (\textit{filter}\ (\lambda \textit{id}.\ \textit{is-RefNode}\ (\textit{kind}\ g\ \textit{id}))\ (\textit{sorted-list-of-set}\ (\textit{ids}\ g))))$

fun $\textit{replace-ref-nodes} :: \textit{IRGraph} \Rightarrow (\textit{ID} \rightharpoonup \textit{ID}) \Rightarrow \textit{ID list} \Rightarrow \textit{ID list}$ **where**
$\textit{replace-ref-nodes}\ g\ m\ \textit{xs} = \textit{map}\ (\lambda \textit{id}.\ (\textit{case}\ (m\ \textit{id})\ \textit{of}\ \textit{Some other} \Rightarrow \textit{other}\ |\ \textit{None} \Rightarrow \textit{id}))\ \textit{xs}$

fun $\textit{find-next} :: \textit{ID list} \Rightarrow \textit{ID set} \Rightarrow \textit{ID option}$ **where**
$\textit{find-next}\ \textit{to-see}\ \textit{seen} = (\textit{let}\ l = (\textit{filter}\ (\lambda \textit{nid}.\ \textit{nid} \notin \textit{seen})\ \textit{to-see})$
$\textit{in}\ (\textit{case}\ l\ \textit{of}\ [] \Rightarrow \textit{None}\ |\ \textit{xs} \Rightarrow \textit{Some}\ (\textit{hd}\ \textit{xs})))$

inductive $\textit{reachables} :: \textit{IRGraph} \Rightarrow \textit{ID list} \Rightarrow \textit{ID set} \Rightarrow \textit{ID set} \Rightarrow \textit{bool}$ **where**
$\textit{reachables}\ g\ []\ \{\}\ \{\}\ |$
$\llbracket \textit{None} = \textit{find-next}\ \textit{to-see}\ \textit{seen}\rrbracket \Longrightarrow \textit{reachables}\ g\ \textit{to-see}\ \textit{seen}\ \textit{seen}\ |$
$\llbracket \textit{Some}\ n = \textit{find-next}\ \textit{to-see}\ \textit{seen};$
$\textit{node} = \textit{kind}\ g\ n;$
$\textit{new} = (\textit{inputs-of}\ \textit{node})\ @\ (\textit{successors-of}\ \textit{node});$
$\textit{reachables}\ g\ (\textit{to-see}\ @\ \textit{new})\ (\{n\} \cup \textit{seen})\ \textit{seen}'\ \rrbracket \Longrightarrow \textit{reachables}\ g\ \textit{to-see}\ \textit{seen}\ \textit{seen}'$

code-pred (*modes*: $i \Rightarrow i \Rightarrow i \Rightarrow o \Rightarrow bool$) [*show-steps*,*show-mode-inference*,*show-intermediate-results*]

reachables ⟨*proof*⟩

inductive *nodeEq* :: $(ID \rightharpoonup ID) \Rightarrow IRGraph \Rightarrow ID \Rightarrow IRGraph \Rightarrow ID \Rightarrow bool$
where
$[\![$ *kind g1 n1* = *RefNode ref*; *nodeEq m g1 ref g2 n2* $]\!] \Longrightarrow$ *nodeEq m g1 n1 g2 n2* |
$[\![$ *x* = *kind g1 n1*;
  y = *kind g2 n2*;
  is-same-ir-node-type x y;
  replace-ref-nodes g1 m (*successors-of x*) = *successors-of y*;
  replace-ref-nodes g1 m (*inputs-of x*) = *inputs-of y* $]\!]$
  $\Longrightarrow$ *nodeEq m g1 n1 g2 n2*

code-pred [*show-modes*] *nodeEq* ⟨*proof*⟩

fun *diffNodesGraph* :: $IRGraph \Rightarrow IRGraph \Rightarrow ID\ set$ **where**
diffNodesGraph g1 g2 = (*let refNodes* = *find-ref-nodes g1 in*
  { *n* . $n \in$ *Predicate.the* (*reachables-i-i-i-o g1* [*0*] {}) $\wedge$ (*case refNodes n of Some* - $\Rightarrow$ *False* | - $\Rightarrow$ *True*) $\wedge$ $\neg$(*nodeEq refNodes g1 n g2 n*)})

fun *diffNodesInfo* :: $IRGraph \Rightarrow IRGraph \Rightarrow (ID \times IRNode \times IRNode)\ set$ (**infix** $\cap_s$ *20*)
  where
diffNodesInfo g1 g2 = {(*nid*, *kind g1 nid*, *kind g2 nid*) | *nid* . $nid \in$ *diffNodesGraph g1 g2*}

fun *eqGraph* :: $IRGraph \Rightarrow IRGraph \Rightarrow bool$ (**infix** $\approx_s$ *20*)
  where
eqGraph isabelle-graph graal-graph = ((*diffNodesGraph isabelle-graph graal-graph*) = {})

end

4 java.lang.Long

Utility functions from the Long class that Graal occasionally makes use of.

theory *Long*
  imports *ValueThms*
begin

lemma *negative-all-set-32*:
  $n < 32 \Longrightarrow bit\ (-1::int32)\ n$
  ⟨*proof*⟩

definition *MaxOrNeg* :: $nat\ set \Rightarrow int$ **where**

$MaxOrNeg\ s = (if\ s = \{\}\ then\ -1\ else\ Max\ s)$

definition $MinOrHighest :: nat\ set \Rightarrow nat \Rightarrow nat$ **where**
$MinOrHighest\ s\ m = (if\ s = \{\}\ then\ m\ else\ Min\ s)$

definition $highestOneBit :: ('a::len)\ word \Rightarrow int$ **where**
$highestOneBit\ v = MaxOrNeg\ \{n\ .\ bit\ v\ n\}$

definition $lowestOneBit :: ('a::len)\ word \Rightarrow nat$ **where**
$lowestOneBit\ v = MinOrHighest\ \{n\ .\ bit\ v\ n\}\ (size\ v)$

lemma *max-bit*: $bit\ (v::('a::len)\ word)\ n \Longrightarrow n < size\ v$
⟨*proof*⟩

lemma *max-set-bit*: $MaxOrNeg\ \{n\ .\ bit\ (v::('a::len)\ word)\ n\} < Nat.size\ v$
⟨*proof*⟩

4.1 Long.numberOfLeadingZeros

definition $numberOfLeadingZeros :: ('a::len)\ word \Rightarrow nat$ **where**
$numberOfLeadingZeros\ v = nat\ (Nat.size\ v - highestOneBit\ v - 1)$

lemma *MaxOrNeg-neg*: $MaxOrNeg\ \{\} = -1$
⟨*proof*⟩

lemma *MaxOrNeg-max*: $s \neq \{\} \Longrightarrow MaxOrNeg\ s = Max\ s$
⟨*proof*⟩

lemma *zero-no-bits*:
$\{n\ .\ bit\ 0\ n\} = \{\}$
⟨*proof*⟩

lemma $highestOneBit\ (0::64\ word) = -1$
⟨*proof*⟩

lemma $numberOfLeadingZeros\ (0::64\ word) = 64$
⟨*proof*⟩

lemma *highestOneBit-top*: $Max\ \{highestOneBit\ (v::64\ word)\} < 64$
⟨*proof*⟩

lemma *numberOfLeadingZeros-top*: $Max\ \{numberOfLeadingZeros\ (v::64\ word)\} \leq 64$
⟨*proof*⟩

lemma *numberOfLeadingZeros-range*: $0 \leq numberOfLeadingZeros\ a \wedge numberOfLeadingZeros\ a \leq Nat.size\ a$
⟨*proof*⟩

lemma *leadingZerosAddHighestOne*: $numberOfLeadingZeros\ v + highestOneBit\ v = Nat.size\ v - 1$
⟨*proof*⟩

4.2 Long.numberOfTrailingZeros

definition *numberOfTrailingZeros* :: $('a::len)\ word \Rightarrow nat$ **where**
$numberOfTrailingZeros\ v = lowestOneBit\ v$

lemma *lowestOneBit-bot*: $lowestOneBit\ (0::64\ word) = 64$
⟨*proof*⟩

lemma *bit-zero-set-in-top*: $bit\ (-1::'a::len\ word)\ 0$
⟨*proof*⟩

lemma *nat-bot-set*: $(0::nat) \in xs \longrightarrow (\forall x \in xs\ .\ 0 \leq x)$
⟨*proof*⟩

lemma $numberOfTrailingZeros\ (0::64\ word) = 64$
⟨*proof*⟩

4.3 Long.bitCount

definition *bitCount* :: $('a::len)\ word \Rightarrow nat$ **where**
$bitCount\ v = card\ \{n\ .\ bit\ v\ n\}$

lemma $bitCount\ 0 = 0$
⟨*proof*⟩

4.4 Long.zeroCount

definition *zeroCount* :: $('a::len)\ word \Rightarrow nat$ **where**
$zeroCount\ v = card\ \{n.\ n < Nat.size\ v \wedge \neg(bit\ v\ n)\}$

lemma *zeroCount-finite*: $finite\ \{n.\ n < Nat.size\ v \wedge \neg(bit\ v\ n)\}$
⟨*proof*⟩

lemma *negone-set*:
$bit\ (-1::('a::len)\ word)\ n \longleftrightarrow n < LENGTH('a)$
⟨*proof*⟩

lemma *negone-all-bits*:
$\{n\ .\ bit\ (-1::('a::len)\ word)\ n\} = \{n\ .\ 0 \leq n \wedge n < LENGTH('a)\}$
⟨*proof*⟩

lemma *bitCount-finite*:
$finite\ \{n\ .\ bit\ (v::('a::len)\ word)\ n\}$
⟨*proof*⟩

lemma *card-of-range*:
$x = card \{n \,.\, 0 \leq n \wedge n < x\}$
⟨*proof*⟩

lemma *range-of-nat*:
$\{(n::nat) \,.\, 0 \leq n \wedge n < x\} = \{n \,.\, n < x\}$
⟨*proof*⟩

lemma *finite-range*:
$finite\ \{n::nat \,.\, n < x\}$
⟨*proof*⟩

lemma *range-eq*:
fixes $x\ y :: nat$
shows $card\ \{y..{<}x\} = card\ \{y{<}..x\}$
⟨*proof*⟩

lemma *card-of-range-bound*:
fixes $x\ y :: nat$
assumes $x > y$
shows $x - y = card\ \{n \,.\, y < n \wedge n \leq x\}$
⟨*proof*⟩

lemma $bitCount\ (-1::('a::len)\ word) = LENGTH('a)$
⟨*proof*⟩

lemma *bitCount-range*:
fixes $n :: ('a::len)\ word$
shows $0 \leq bitCount\ n \wedge bitCount\ n \leq Nat.size\ n$
⟨*proof*⟩

lemma *zerosAboveHighestOne*:
$n > highestOneBit\ a \Longrightarrow \neg(bit\ a\ n)$
⟨*proof*⟩

lemma *zerosBelowLowestOne*:
assumes $n < lowestOneBit\ a$
shows $\neg(bit\ a\ n)$
⟨*proof*⟩

lemma *union-bit-sets*:
fixes $a :: ('a::len)\ word$
shows $\{n \,.\, n < Nat.size\ a \wedge bit\ a\ n\} \cup \{n \,.\, n < Nat.size\ a \wedge \neg(bit\ a\ n)\} = \{n \,.\, n < Nat.size\ a\}$
⟨*proof*⟩

lemma *disjoint-bit-sets*:
fixes $a :: ('a::len)\ word$

shows $\{n\ .\ n < Nat.size\ a \wedge bit\ a\ n\} \cap \{n\ .\ n < Nat.size\ a \wedge \neg(bit\ a\ n)\} = \{\}$
⟨*proof*⟩

lemma *qualified-bitCount*:
bitCount v = card $\{n\ .\ n < Nat.size\ v \wedge bit\ v\ n\}$
⟨*proof*⟩

lemma *card-eq*:
assumes *finite x* ∧ *finite y* ∧ *finite z*
assumes $x \cup y = z$
assumes $y \cap x = \{\}$
shows *card z* − *card y* = *card x*
⟨*proof*⟩

lemma *card-add*:
assumes *finite x* ∧ *finite y* ∧ *finite z*
assumes $x \cup y = z$
assumes $y \cap x = \{\}$
shows *card x* + *card y* = *card z*
⟨*proof*⟩

lemma *card-add-inverses*:
assumes *finite* $\{n.\ Q\ n \wedge \neg(P\ n)\}$ ∧ *finite* $\{n.\ Q\ n \wedge P\ n\}$ ∧ *finite* $\{n.\ Q\ n\}$
shows *card* $\{n.\ Q\ n \wedge P\ n\}$ + *card* $\{n.\ Q\ n \wedge \neg(P\ n)\}$ = *card* $\{n.\ Q\ n\}$
⟨*proof*⟩

lemma *ones-zero-sum-to-width*:
bitCount a + *zeroCount a* = *Nat.size a*
⟨*proof*⟩

lemma *intersect-bitCount-helper*:
card $\{n\ .\ n < Nat.size\ a\}$ − *bitCount a* = *card* $\{n\ .\ n < Nat.size\ a \wedge \neg(bit\ a\ n)\}$
⟨*proof*⟩

lemma *intersect-bitCount*:
Nat.size a − *bitCount a* = *card* $\{n\ .\ n < Nat.size\ a \wedge \neg(bit\ a\ n)\}$
⟨*proof*⟩

hide-fact *intersect-bitCount-helper*

end

5 Data-flow Semantics

theory *IRTreeEval*
imports
Graph.Stamp
begin

We define a tree representation of data-flow nodes, as an abstraction of the graph view.

Data-flow trees are evaluated in the context of a method state (currently called MapState in the theories for historical reasons).

The method state consists of the values for each method parameter, references to method parameters use an index of the parameter within the parameter list, as such we store a list of parameter values which are looked up at parameter references.

The method state also stores a mapping of node ids to values. The contents of this mapping is calculates during the traversal of the control flow graph.

As a concrete example, as the *SignedDivNode* can have side-effects (during division by zero), it is treated as part of the control-flow, since the data-flow phase is specified to be side-effect free. As a result, the control-flow semantics for *SignedDivNode* calculates the value of a node and maps the node identifier to the value within the method state. The data-flow semantics then just reads the value stored in the method state for the node.

type-synonym *ID* = *nat*
type-synonym *MapState* = *ID* ⇒ *Value*
type-synonym *Params* = *Value list*

definition *new-map-state* :: *MapState* **where**
  new-map-state = (λ*x*. *UndefVal*)

5.1 Data-flow Tree Representation

datatype *IRUnaryOp* =
    UnaryAbs
  | *UnaryNeg*
  | *UnaryNot*
  | *UnaryLogicNegation*
  | *UnaryNarrow* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*)
  | *UnarySignExtend* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*)
  | *UnaryZeroExtend* (*ir-inputBits*: *nat*) (*ir-resultBits*: *nat*)

datatype *IRBinaryOp* =
    BinAdd
  | *BinMul*
  | *BinSub*
  | *BinAnd*
  | *BinOr*
  | *BinXor*
  | *BinShortCircuitOr*
  | *BinLeftShift*
  | *BinRightShift*
  | *BinURightShift*
  | *BinIntegerEquals*

| *BinIntegerLessThan*
| *BinIntegerBelow*

datatype (*discs-sels*) *IRExpr* =
UnaryExpr (*ir-uop*: *IRUnaryOp*) (*ir-value*: *IRExpr*)
| *BinaryExpr* (*ir-op*: *IRBinaryOp*) (*ir-x*: *IRExpr*) (*ir-y*: *IRExpr*)
| *ConditionalExpr* (*ir-condition*: *IRExpr*) (*ir-trueValue*: *IRExpr*) (*ir-falseValue*: *IRExpr*)

| *ParameterExpr* (*ir-index*: *nat*) (*ir-stamp*: *Stamp*)

| *LeafExpr* (*ir-nid*: *ID*) (*ir-stamp*: *Stamp*)

| *ConstantExpr* (*ir-const*: *Value*)
| *ConstantVar* (*ir-name*: *string*)
| *VariableExpr* (*ir-name*: *string*) (*ir-stamp*: *Stamp*)

fun *is-ground* :: *IRExpr* ⇒ *bool* **where**
is-ground (*UnaryExpr op e*) = *is-ground e* |
is-ground (*BinaryExpr op e1 e2*) = (*is-ground e1* ∧ *is-ground e2*) |
is-ground (*ConditionalExpr b e1 e2*) = (*is-ground b* ∧ *is-ground e1* ∧ *is-ground e2*) |
is-ground (*ParameterExpr i s*) = *True* |
is-ground (*LeafExpr n s*) = *True* |
is-ground (*ConstantExpr v*) = *True* |
is-ground (*ConstantVar name*) = *False* |
is-ground (*VariableExpr name s*) = *False*

typedef *GroundExpr* = { *e* :: *IRExpr* . *is-ground e* }
⟨*proof*⟩

5.2 Functions for re-calculating stamps

Note: in Java all integer calculations are done as 32 or 64 bit calculations. However, here we generalise the operators to allow any size calculations. Many operators have the same output bits as their inputs. However, the unary integer operators that are not *normal_unary* are narrowing or widening operators, so the result bits is specified by the operator. The binary integer operators are divided into three groups: (1) *binary_fixed_*32 operators always output 32 bits, (2) *binary_shift_ops* operators output size is determined by their left argument, and (3) other operators output the same number of bits as both their inputs.

abbreviation *binary-fixed-32-ops* :: *IRBinaryOp set* **where**
binary-fixed-32-ops ≡ {*BinShortCircuitOr*, *BinIntegerEquals*, *BinIntegerLessThan*, *BinIntegerBelow*}

abbreviation *binary-shift-ops* :: *IRBinaryOp set* **where**

binary-shift-ops ≡ {*BinLeftShift*, *BinRightShift*, *BinURightShift*}

abbreviation *normal-unary* :: *IRUnaryOp set* **where**
normal-unary ≡ {*UnaryAbs*, *UnaryNeg*, *UnaryNot*, *UnaryLogicNegation*}

fun *stamp-unary* :: *IRUnaryOp* ⇒ *Stamp* ⇒ *Stamp* **where**

stamp-unary op (*IntegerStamp b lo hi*) =
unrestricted-stamp (*IntegerStamp* (*if op* ∈ *normal-unary then b else* (*ir-resultBits op*)) *lo hi*) |

stamp-unary op - = *IllegalStamp*

fun *stamp-binary* :: *IRBinaryOp* ⇒ *Stamp* ⇒ *Stamp* ⇒ *Stamp* **where**
stamp-binary op (*IntegerStamp b1 lo1 hi1*) (*IntegerStamp b2 lo2 hi2*) =
(*if op* ∈ *binary-shift-ops then unrestricted-stamp* (*IntegerStamp b1 lo1 hi1*)
else if b1 ≠ *b2 then IllegalStamp else*
(*if op* ∈ *binary-fixed-32-ops*
then unrestricted-stamp (*IntegerStamp 32 lo1 hi1*)
else unrestricted-stamp (*IntegerStamp b1 lo1 hi1*))) |

stamp-binary op - - = *IllegalStamp*

fun *stamp-expr* :: *IRExpr* ⇒ *Stamp* **where**
stamp-expr (*UnaryExpr op x*) = *stamp-unary op* (*stamp-expr x*) |
stamp-expr (*BinaryExpr bop x y*) = *stamp-binary bop* (*stamp-expr x*) (*stamp-expr y*) |
stamp-expr (*ConstantExpr val*) = *constantAsStamp val* |
stamp-expr (*LeafExpr i s*) = *s* |
stamp-expr (*ParameterExpr i s*) = *s* |
stamp-expr (*ConditionalExpr c t f*) = *meet* (*stamp-expr t*) (*stamp-expr f*)

export-code *stamp-unary stamp-binary stamp-expr*

5.3 Data-flow Tree Evaluation

fun *unary-eval* :: *IRUnaryOp* ⇒ *Value* ⇒ *Value* **where**
unary-eval UnaryAbs v = *intval-abs v* |
unary-eval UnaryNeg v = *intval-negate v* |
unary-eval UnaryNot v = *intval-not v* |
unary-eval UnaryLogicNegation v = *intval-logic-negation v* |
unary-eval (*UnaryNarrow inBits outBits*) *v* = *intval-narrow inBits outBits v* |
unary-eval (*UnarySignExtend inBits outBits*) *v* = *intval-sign-extend inBits outBits v* |
unary-eval (*UnaryZeroExtend inBits outBits*) *v* = *intval-zero-extend inBits outBits v*

fun *bin-eval* :: *IRBinaryOp* ⇒ *Value* ⇒ *Value* ⇒ *Value* **where**

bin-eval BinAdd v1 v2 = *intval-add v1 v2* |
bin-eval BinMul v1 v2 = *intval-mul v1 v2* |
bin-eval BinSub v1 v2 = *intval-sub v1 v2* |
bin-eval BinAnd v1 v2 = *intval-and v1 v2* |
bin-eval BinOr v1 v2 = *intval-or v1 v2* |
bin-eval BinXor v1 v2 = *intval-xor v1 v2* |
bin-eval BinShortCircuitOr v1 v2 = *intval-short-circuit-or v1 v2* |
bin-eval BinLeftShift v1 v2 = *intval-left-shift v1 v2* |
bin-eval BinRightShift v1 v2 = *intval-right-shift v1 v2* |
bin-eval BinURightShift v1 v2 = *intval-uright-shift v1 v2* |
bin-eval BinIntegerEquals v1 v2 = *intval-equals v1 v2* |
bin-eval BinIntegerLessThan v1 v2 = *intval-less-than v1 v2* |
bin-eval BinIntegerBelow v1 v2 = *intval-below v1 v2*

lemmas *eval-thms* =
intval-abs.simps intval-negate.simps intval-not.simps
intval-logic-negation.simps intval-narrow.simps
intval-sign-extend.simps intval-zero-extend.simps
intval-add.simps intval-mul.simps intval-sub.simps
intval-and.simps intval-or.simps intval-xor.simps
intval-left-shift.simps intval-right-shift.simps
intval-uright-shift.simps intval-equals.simps
intval-less-than.simps intval-below.simps

inductive *not-undef-or-fail* :: *Value* ⇒ *Value* ⇒ *bool* **where**
⟦*value* ≠ *UndefVal*⟧ ⟹ *not-undef-or-fail value value*

notation (*latex* **output**)
not-undef-or-fail (- = -)

inductive
evaltree :: *MapState* ⇒ *Params* ⇒ *IRExpr* ⇒ *Value* ⇒ *bool* ([-,-] ⊢ - ↦ - *55*)
for *m p* **where**

ConstantExpr:
⟦*wf-value c*⟧
⟹ [*m*,*p*] ⊢ (*ConstantExpr c*) ↦ *c* |

ParameterExpr:
⟦*i* < *length p*; *valid-value* (*p*!*i*) *s*⟧
⟹ [*m*,*p*] ⊢ (*ParameterExpr i s*) ↦ *p*!*i* |

ConditionalExpr:
⟦[*m*,*p*] ⊢ *ce* ↦ *cond*;
cond ≠ *UndefVal*;
branch = (*if val-to-bool cond then te else fe*);
[*m*,*p*] ⊢ *branch* ↦ *result*;

result ≠ *UndefVal*⟧
  ⟹ [*m*,*p*] ⊢ (*ConditionalExpr ce te fe*) ↦ *result* |

UnaryExpr:
⟦[*m*,*p*] ⊢ *xe* ↦ *x*;
  result = (*unary-eval op x*);
  result ≠ *UndefVal*⟧
  ⟹ [*m*,*p*] ⊢ (*UnaryExpr op xe*) ↦ *result* |

BinaryExpr:
⟦[*m*,*p*] ⊢ *xe* ↦ *x*;
  [*m*,*p*] ⊢ *ye* ↦ *y*;
  result = (*bin-eval op x y*);
  result ≠ *UndefVal*⟧
  ⟹ [*m*,*p*] ⊢ (*BinaryExpr op xe ye*) ↦ *result* |

LeafExpr:
⟦*val* = *m n*;
  valid-value val s⟧
  ⟹ [*m*,*p*] ⊢ *LeafExpr n s* ↦ *val*

code-pred (*modes*: *i* ⇒ *i* ⇒ *i* ⇒ *o* ⇒ *bool as evalT*)
  [*show-steps*,*show-mode-inference*,*show-intermediate-results*]
  evaltree ⟨*proof*⟩

inductive
  evaltrees :: *MapState* ⇒ *Params* ⇒ *IRExpr list* ⇒ *Value list* ⇒ *bool* ([-,-] ⊢ - $\mapsto_L$ - *55*)
  for *m p* **where**

EvalNil:
[*m*,*p*] ⊢ [] $\mapsto_L$ [] |

EvalCons:
⟦[*m*,*p*] ⊢ *x* ↦ *xval*;
  [*m*,*p*] ⊢ *yy* $\mapsto_L$ *yyval*⟧
  ⟹ [*m*,*p*] ⊢ (*x*#*yy*) $\mapsto_L$ (*xval*#*yyval*)

code-pred (*modes*: *i* ⇒ *i* ⇒ *i* ⇒ *o* ⇒ *bool as evalTs*)
  evaltrees ⟨*proof*⟩

definition *sq-param0* :: *IRExpr* **where**
  sq-param0 = *BinaryExpr BinMul*
    (*ParameterExpr 0* (*IntegerStamp 32* (− *2147483648*) *2147483647*))
    (*ParameterExpr 0* (*IntegerStamp 32* (− *2147483648*) *2147483647*))

values {*v*. *evaltree new-map-state* [*IntVal 32 5*] *sq-param0 v*}

declare *evaltree.intros* [*intro*]
declare *evaltrees.intros* [*intro*]

5.4 Data-flow Tree Refinement

We define the induced semantic equivalence relation between expressions. Note that syntactic equality implies semantic equivalence, but not vice versa.

definition *equiv-exprs* :: *IRExpr* $\Rightarrow$ *IRExpr* $\Rightarrow$ *bool* (- $\doteq$ - *55*) **where**
$(e1 \doteq e2) = (\forall\ m\ p\ v.\ (([m,p] \vdash e1 \mapsto v) \longleftrightarrow ([m,p] \vdash e2 \mapsto v)))$

We also prove that this is a total equivalence relation (*equivp equiv-exprs*) (HOL.Equiv_Relations), so that we can reuse standard results about equivalence relations.

lemma *equivp equiv-exprs*
⟨*proof*⟩

We define a refinement ordering over IRExpr and show that it is a preorder. Note that it is asymmetric because e2 may refer to fewer variables than e1.

instantiation *IRExpr* :: *preorder* **begin**

notation *less-eq* (**infix** $\sqsubseteq$ *65*)

definition
le-expr-def [*simp*]:
$(e_2 \leq e_1) \longleftrightarrow (\forall\ m\ p\ v.\ (([m,p] \vdash e_1 \mapsto v) \longrightarrow ([m,p] \vdash e_2 \mapsto v)))$

definition
lt-expr-def [*simp*]:
$(e_1 < e_2) \longleftrightarrow (e_1 \leq e_2 \wedge \neg\ (e_1 \doteq e_2))$

instance ⟨*proof*⟩

end

abbreviation (**output**) *Refines* :: *IRExpr* $\Rightarrow$ *IRExpr* $\Rightarrow$ *bool* (**infix** $\sqsupseteq$ *64*)
where $e_1 \sqsupseteq e_2 \equiv (e_2 \leq e_1)$

5.5 Stamp Masks

A stamp can contain additional range information in the form of masks. A stamp has an up mask and a down mask, corresponding to a the bits that may be set and the bits that must be set.

Examples: A stamp where no range information is known will have; an up mask of -1 as all bits may be set, and a down mask of 0 as no bits must be set.

A stamp known to be one should have; an up mask of 1 as only the first bit may be set, no others, and a down mask of 1 as the first bit must be set and no others.

We currently don't carry mask information in stamps, and instead assume correct masks to prove optimizations.

locale *stamp-mask* =
  fixes *up* :: *IRExpr* ⇒ *int64* (↑)
  fixes *down* :: *IRExpr* ⇒ *int64* (↓)
  assumes *up-spec*: $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow (and\ v\ (not\ ((ucast\ (\uparrow e))))) = 0$
    and *down-spec*: $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow (and\ (not\ v)\ (ucast\ (\downarrow e))) = 0$
begin

lemma *may-implies-either*:
  $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow bit\ (\uparrow e)\ n \Longrightarrow bit\ v\ n = False \lor bit\ v\ n = True$
  ⟨*proof*⟩

lemma *not-may-implies-false*:
  $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow \neg(bit\ (\uparrow e)\ n) \Longrightarrow bit\ v\ n = False$
  ⟨*proof*⟩

lemma *must-implies-true*:
  $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow bit\ (\downarrow e)\ n \Longrightarrow bit\ v\ n = True$
  ⟨*proof*⟩

lemma *not-must-implies-either*:
  $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow \neg(bit\ (\downarrow e)\ n) \Longrightarrow bit\ v\ n = False \lor bit\ v\ n = True$
  ⟨*proof*⟩

lemma *must-implies-may*:
  $[m, p] \vdash e \mapsto IntVal\ b\ v \Longrightarrow n < 32 \Longrightarrow bit\ (\downarrow e)\ n \Longrightarrow bit\ (\uparrow e)\ n$
  ⟨*proof*⟩

lemma *up-mask-and-zero-implies-zero*:
  assumes $and\ (\uparrow x)\ (\uparrow y) = 0$
  assumes $[m, p] \vdash x \mapsto IntVal\ b\ xv$
  assumes $[m, p] \vdash y \mapsto IntVal\ b\ yv$
  shows $and\ xv\ yv = 0$
  ⟨*proof*⟩

lemma *not-down-up-mask-and-zero-implies-zero*:
  assumes $and\ (not\ (\downarrow x))\ (\uparrow y) = 0$
  assumes $[m, p] \vdash x \mapsto IntVal\ b\ xv$
  assumes $[m, p] \vdash y \mapsto IntVal\ b\ yv$
  shows $and\ xv\ yv = yv$
  ⟨*proof*⟩

end

definition *IRExpr-up* :: *IRExpr* $\Rightarrow$ *int64* **where**
IRExpr-up e = *not 0*

definition *IRExpr-down* :: *IRExpr* $\Rightarrow$ *int64* **where**
IRExpr-down e = *0*

lemma *ucast-zero*: $(ucast\ (0::int64)::int32) = 0$
⟨*proof*⟩

lemma *ucast-minus-one*: $(ucast\ (-1::int64)::int32) = -1$
⟨*proof*⟩

interpretation *simple-mask*: *stamp-mask*
IRExpr-up :: *IRExpr* $\Rightarrow$ *int64*
IRExpr-down :: *IRExpr* $\Rightarrow$ *int64*
⟨*proof*⟩

end

5.6 Data-flow Tree Theorems

theory *IRTreeEvalThms*
imports
Graph.ValueThms
IRTreeEval
begin

5.6.1 Deterministic Data-flow Evaluation

lemma *evalDet*:
$[m,p] \vdash e \mapsto v_1 \Longrightarrow$
$[m,p] \vdash e \mapsto v_2 \Longrightarrow$
$v_1 = v_2$
⟨*proof*⟩

lemma *evalAllDet*:
$[m,p] \vdash e \mapsto_L v1 \Longrightarrow$
$[m,p] \vdash e \mapsto_L v2 \Longrightarrow$
$v1 = v2$
⟨*proof*⟩

5.6.2 Typing Properties for Integer Evaluation Functions

We use three simple typing properties on integer values: $is_IntVal32, is_IntVal64$ and the more general is_IntVal.

lemma *unary-eval-not-obj-ref*:
shows *unary-eval op x* $\neq$ *ObjRef v*
⟨*proof*⟩

lemma *unary-eval-not-obj-str*:
  shows $\textit{unary-eval op } x \neq \textit{ObjStr } v$
  ⟨*proof*⟩

lemma *unary-eval-int*:
  assumes *def*: $\textit{unary-eval op } x \neq \textit{UndefVal}$
  shows $\textit{is-IntVal } (\textit{unary-eval op } x)$
  ⟨*proof*⟩

lemma *bin-eval-int*:
  assumes *def*: $\textit{bin-eval op } x\ y \neq \textit{UndefVal}$
  shows $\textit{is-IntVal } (\textit{bin-eval op } x\ y)$
  ⟨*proof*⟩

lemma *IntVal0*:
  $(\textit{IntVal } 32\ 0) = (\textit{new-int } 32\ 0)$
  ⟨*proof*⟩

lemma *IntVal1*:
  $(\textit{IntVal } 32\ 1) = (\textit{new-int } 32\ 1)$
  ⟨*proof*⟩

lemma *bin-eval-new-int*:
  assumes *def*: $\textit{bin-eval op } x\ y \neq \textit{UndefVal}$
  shows $\exists\, b\ v.\ (\textit{bin-eval op } x\ y) = \textit{new-int } b\ v\ \wedge$
  $b = (\textit{if op} \in \textit{binary-fixed-32-ops then } 32 \textit{ else intval-bits } x)$
  ⟨*proof*⟩

lemma *int-stamp*:
  assumes *i*: $\textit{is-IntVal } v$
  shows $\textit{is-IntegerStamp } (\textit{constantAsStamp } v)$
  ⟨*proof*⟩

lemma *validStampIntConst*:
  assumes $v = \textit{IntVal } b\ \textit{ival}$
  assumes $0 < b \wedge b \leq 64$
  shows $\textit{valid-stamp } (\textit{constantAsStamp } v)$
⟨*proof*⟩

lemma *validDefIntConst*:
assumes *v*: $v = IntVal\ b\ ival$
assumes $0 < b \wedge b \leq 64$
assumes *take-bit b ival = ival*
shows *valid-value v* (*constantAsStamp v*)
⟨*proof*⟩

5.6.3 Evaluation Results are Valid

A valid value cannot be $UndefVal$.

lemma *valid-not-undef*:
assumes *a1*: *valid-value val s*
assumes *a2*: $s \neq VoidStamp$
shows $val \neq UndefVal$
⟨*proof*⟩

lemma *valid-VoidStamp*[*elim*]:
shows *valid-value val VoidStamp* $\Longrightarrow$
$val = UndefVal$
⟨*proof*⟩

lemma *valid-ObjStamp*[*elim*]:
shows *valid-value val* (*ObjectStamp klass exact nonNull alwaysNull*) $\Longrightarrow$
($\exists\, v.\ val = ObjRef\ v$)
⟨*proof*⟩

lemma *valid-int*[*elim*]:
shows *valid-value val* (*IntegerStamp b lo hi*) $\Longrightarrow$
($\exists\, v.\ val = IntVal\ b\ v$)
⟨*proof*⟩

lemmas *valid-value-elims* =
valid-VoidStamp
valid-ObjStamp
valid-int

lemma *evaltree-not-undef*:
fixes *m p e v*
shows ($[m,p] \vdash e \mapsto v$) $\Longrightarrow v \neq UndefVal$
⟨*proof*⟩

lemma *leafint*:
assumes *ev*: $[m,p] \vdash LeafExpr\ i\ (IntegerStamp\ b\ lo\ hi) \mapsto val$
shows $\exists\, b\ v.\ val = (IntVal\ b\ v)$

⟨*proof*⟩

lemma *default-stamp* [*simp*]: *default-stamp* = *IntegerStamp* *32* (−*2147483648*) *2147483647*
⟨*proof*⟩

lemma *valid-value-signed-int-range* [*simp*]:
assumes *valid-value val* (*IntegerStamp b lo hi*)
assumes $lo < 0$
shows $\exists v.\ (val = IntVal\ b\ v \wedge$
$lo \leq int\text{-}signed\text{-}value\ b\ v \wedge$
$int\text{-}signed\text{-}value\ b\ v \leq hi)$
⟨*proof*⟩

5.6.4 Example Data-flow Optimisations

5.6.5 Monotonicity of Expression Refinement

We prove that each subexpression position is monotonic. That is, optimizing a subexpression anywhere deep inside a top-level expression also optimizes that top-level expression.

Note that we might also be able to do this via reusing Isabelle's *mono* operator (HOL.Orderings theory), proving instantiations like $mono(UnaryExprop)$, but it is not obvious how to do this for both arguments of the binary expressions.

lemma *mono-unary*:
assumes $x \geq x'$
shows $(UnaryExpr\ op\ x) \geq (UnaryExpr\ op\ x')$
⟨*proof*⟩

lemma *mono-binary*:
assumes $x \geq x'$
assumes $y \geq y'$
shows $(BinaryExpr\ op\ x\ y) \geq (BinaryExpr\ op\ x'\ y')$
⟨*proof*⟩

lemma *never-void*:
assumes $[m, p] \vdash x \mapsto xv$
assumes *valid-value xv* (*stamp-expr xe*)
shows $stamp\text{-}expr\ xe \neq VoidStamp$
⟨*proof*⟩

lemma *compatible-trans*:
$compatible\ x\ y \wedge compatible\ y\ z \Longrightarrow compatible\ x\ z$

$\langle proof \rangle$

lemma *compatible-refl*:
  compatible x y $\Longrightarrow$ *compatible y x*
  $\langle proof \rangle$

lemma *mono-conditional*:
  assumes $c \geq c'$
  assumes $t \geq t'$
  assumes $f \geq f'$
  shows $(ConditionalExpr\ c\ t\ f) \geq (ConditionalExpr\ c'\ t'\ f')$
$\langle proof \rangle$

5.7 Unfolding rules for evaltree quadruples down to bin-eval level

These rewrite rules can be useful when proving optimizations. They support top-down rewriting of each level of the tree into the lower-level $bin_e val$ / $unary_e val$ level, simply by saying $unfoldingunfold_e valtree$.

lemma *unfold-const*:
  shows $([m,p] \vdash ConstantExpr\ c \mapsto v) = (wf\text{-}value\ v \wedge v = c)$
  $\langle proof \rangle$

lemma *unfold-binary*:
  shows $([m,p] \vdash BinaryExpr\ op\ xe\ ye \mapsto val) = (\exists\ x\ y.$
    $(([m,p] \vdash xe \mapsto x) \wedge$
    $([m,p] \vdash ye \mapsto y) \wedge$
    $(val = bin\text{-}eval\ op\ x\ y) \wedge$
    $(val \neq UndefVal)$
  $))$ (**is** $?L = ?R$)
$\langle proof \rangle$

lemma *unfold-unary*:
  shows $([m,p] \vdash UnaryExpr\ op\ xe \mapsto val)$
    $= (\exists\ x.$
      $(([m,p] \vdash xe \mapsto x) \wedge$
      $(val = unary\text{-}eval\ op\ x) \wedge$
      $(val \neq UndefVal)$
    $))$ (**is** $?L = ?R$)
  $\langle proof \rangle$

lemmas *unfold-evaltree* =

unfold-binary
unfold-unary

5.8 Lemmas about *new_int* and integer eval results.

lemma *unary-eval-new-int*:
assumes *def*: *unary-eval op x* $\neq$ *UndefVal*
shows $\exists$ *b v. unary-eval op x = new-int b v* $\wedge$
b = (if op $\in$ *normal-unary then intval-bits x else ir-resultBits op)*
⟨*proof*⟩

lemma *new-int-unused-bits-zero*:
assumes *IntVal b ival = new-int b ival0*
shows *take-bit b ival = ival*
⟨*proof*⟩

lemma *unary-eval-unused-bits-zero*:
assumes *unary-eval op x = IntVal b ival*
shows *take-bit b ival = ival*
⟨*proof*⟩

lemma *bin-eval-unused-bits-zero*:
assumes *bin-eval op x y = (IntVal b ival)*
shows *take-bit b ival = ival*
⟨*proof*⟩

lemma *eval-unused-bits-zero*:
[m,p] $\vdash$ *xe* $\mapsto$ *(IntVal b ix)* $\Longrightarrow$ *take-bit b ix = ix*
⟨*proof*⟩

lemma *unary-normal-bitsize*:
assumes *unary-eval op x = IntVal b ival*
assumes *op* $\in$ *normal-unary*
shows $\exists$ *ix. x = IntVal b ix*
⟨*proof*⟩

lemma *unary-not-normal-bitsize*:
assumes *unary-eval op x = IntVal b ival*
assumes *op* $\notin$ *normal-unary*
shows *b = ir-resultBits op* $\wedge$ *0 < b* $\wedge$ *b* $\leq$ *64*
⟨*proof*⟩

lemma *unary-eval-bitsize*:
assumes *unary-eval op x = IntVal b ival*
assumes *2*: *x = IntVal bx ix*
assumes *0 < bx* $\wedge$ *bx* $\leq$ *64*
shows *0 < b* $\wedge$ *b* $\leq$ *64*

$\langle proof \rangle$

lemma *bin-eval-inputs-are-ints*:
  assumes *bin-eval op x y = IntVal b ix*
  obtains *xb yb xi yi* **where** $x = IntVal\ xb\ xi \wedge y = IntVal\ yb\ yi$
$\langle proof \rangle$

lemma *eval-bits-1-64*:
  $[m,p] \vdash xe \mapsto (IntVal\ b\ ix) \Longrightarrow 0 < b \wedge b \leq 64$
$\langle proof \rangle$

lemma *unfold-binary-width*:
  assumes $op \notin$ *binary-fixed-32-ops* $\wedge$ $op \notin$ *binary-shift-ops*
  shows $([m,p] \vdash BinaryExpr\ op\ xe\ ye \mapsto IntVal\ b\ val) = (\exists\ x\ y.$
      $(([m,p] \vdash xe \mapsto IntVal\ b\ x) \wedge$
       $([m,p] \vdash ye \mapsto IntVal\ b\ y) \wedge$
       $(IntVal\ b\ val = bin\text{-}eval\ op\ (IntVal\ b\ x)\ (IntVal\ b\ y)) \wedge$
       $(IntVal\ b\ val \neq UndefVal)$
    $))$ (**is** $?L = ?R$)
$\langle proof \rangle$

end

6 Tree to Graph

theory *TreeToGraph*
  imports
    Semantics.IRTreeEval
    Graph.IRGraph
begin

6.1 Subgraph to Data-flow Tree

fun *find-node-and-stamp* :: *IRGraph* $\Rightarrow$ (*IRNode* $\times$ *Stamp*) $\Rightarrow$ *ID option* **where**
  find-node-and-stamp g (*n*,*s*) =
    find ($\lambda i.$ *kind g i = n* $\wedge$ *stamp g i = s*) (*sorted-list-of-set*(*ids g*))

export-code *find-node-and-stamp*

fun *is-preevaluated* :: *IRNode* $\Rightarrow$ *bool* **where**
  is-preevaluated (*InvokeNode n - - - - -*) = *True* |
  is-preevaluated (*InvokeWithExceptionNode n - - - - - -*) = *True* |
  is-preevaluated (*NewInstanceNode n - - -*) = *True* |
  is-preevaluated (*LoadFieldNode n - - -*) = *True* |

is-preevaluated (*SignedDivNode n* - - - - -) = *True* |
is-preevaluated (*SignedRemNode n* - - - - -) = *True* |
is-preevaluated (*ValuePhiNode n* - -) = *True* |
is-preevaluated - = *False*

inductive
rep :: *IRGraph* ⇒ *ID* ⇒ *IRExpr* ⇒ *bool* (- ⊢ - ≃ - *55*)
for *g* **where**

ConstantNode:
⟦*kind g n* = *ConstantNode c*⟧
⟹ *g* ⊢ *n* ≃ (*ConstantExpr c*) |

ParameterNode:
⟦*kind g n* = *ParameterNode i*;
stamp g n = *s*⟧
⟹ *g* ⊢ *n* ≃ (*ParameterExpr i s*) |

ConditionalNode:
⟦*kind g n* = *ConditionalNode c t f*;
g ⊢ *c* ≃ *ce*;
g ⊢ *t* ≃ *te*;
g ⊢ *f* ≃ *fe*⟧
⟹ *g* ⊢ *n* ≃ (*ConditionalExpr ce te fe*) |

AbsNode:
⟦*kind g n* = *AbsNode x*;
g ⊢ *x* ≃ *xe*⟧
⟹ *g* ⊢ *n* ≃ (*UnaryExpr UnaryAbs xe*) |

NotNode:
⟦*kind g n* = *NotNode x*;
g ⊢ *x* ≃ *xe*⟧
⟹ *g* ⊢ *n* ≃ (*UnaryExpr UnaryNot xe*) |

NegateNode:
⟦*kind g n* = *NegateNode x*;
g ⊢ *x* ≃ *xe*⟧
⟹ *g* ⊢ *n* ≃ (*UnaryExpr UnaryNeg xe*) |

LogicNegationNode:
⟦*kind g n* = *LogicNegationNode x*;
g ⊢ *x* ≃ *xe*⟧
⟹ *g* ⊢ *n* ≃ (*UnaryExpr UnaryLogicNegation xe*) |

AddNode:

$[\![kind\ g\ n = AddNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinAdd\ xe\ ye)\ |$

MulNode:
$[\![kind\ g\ n = MulNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinMul\ xe\ ye)\ |$

SubNode:
$[\![kind\ g\ n = SubNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinSub\ xe\ ye)\ |$

AndNode:
$[\![kind\ g\ n = AndNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinAnd\ xe\ ye)\ |$

OrNode:
$[\![kind\ g\ n = OrNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinOr\ xe\ ye)\ |$

XorNode:
$[\![kind\ g\ n = XorNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinXor\ xe\ ye)\ |$

ShortCircuitOrNode:
$[\![kind\ g\ n = ShortCircuitOrNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinShortCircuitOr\ xe\ ye)\ |$

LeftShiftNode:
$[\![kind\ g\ n = LeftShiftNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye]\!]$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinLeftShift\ xe\ ye)\ |$

RightShiftNode:
$[\![kind\ g\ n = RightShiftNode\ x\ y;$

$g \vdash x \simeq xe;$
$g \vdash y \simeq ye\rrbracket$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinRightShift\ xe\ ye) \mid$

UnsignedRightShiftNode:
$\llbracket kind\ g\ n = UnsignedRightShiftNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye\rrbracket$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinURightShift\ xe\ ye) \mid$

IntegerBelowNode:
$\llbracket kind\ g\ n = IntegerBelowNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye\rrbracket$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinIntegerBelow\ xe\ ye) \mid$

IntegerEqualsNode:
$\llbracket kind\ g\ n = IntegerEqualsNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye\rrbracket$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinIntegerEquals\ xe\ ye) \mid$

IntegerLessThanNode:
$\llbracket kind\ g\ n = IntegerLessThanNode\ x\ y;$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye\rrbracket$
$\Longrightarrow g \vdash n \simeq (BinaryExpr\ BinIntegerLessThan\ xe\ ye) \mid$

NarrowNode:
$\llbracket kind\ g\ n = NarrowNode\ inputBits\ resultBits\ x;$
$g \vdash x \simeq xe\rrbracket$
$\Longrightarrow g \vdash n \simeq (UnaryExpr\ (UnaryNarrow\ inputBits\ resultBits)\ xe) \mid$

SignExtendNode:
$\llbracket kind\ g\ n = SignExtendNode\ inputBits\ resultBits\ x;$
$g \vdash x \simeq xe\rrbracket$
$\Longrightarrow g \vdash n \simeq (UnaryExpr\ (UnarySignExtend\ inputBits\ resultBits)\ xe) \mid$

ZeroExtendNode:
$\llbracket kind\ g\ n = ZeroExtendNode\ inputBits\ resultBits\ x;$
$g \vdash x \simeq xe\rrbracket$
$\Longrightarrow g \vdash n \simeq (UnaryExpr\ (UnaryZeroExtend\ inputBits\ resultBits)\ xe) \mid$

LeafNode:
$\llbracket is\text{-}preevaluated\ (kind\ g\ n);$
$stamp\ g\ n = s\rrbracket$
$\Longrightarrow g \vdash n \simeq (LeafExpr\ n\ s) \mid$

RefNode:
$[\![kind\ g\ n = RefNode\ n';$
$\quad g \vdash n' \simeq e]\!]$
$\quad \Longrightarrow g \vdash n \simeq e$

code-pred $(modes: i \Rightarrow i \Rightarrow o \Rightarrow bool\ as\ exprE)\ rep\ \langle proof \rangle$

inductive
$replist :: IRGraph \Rightarrow ID\ list \Rightarrow IRExpr\ list \Rightarrow bool\ (\text{-} \vdash \text{-} \simeq_L \text{-}\ 55)$
for g **where**

RepNil:
$g \vdash [] \simeq_L [] \mid$

RepCons:
$[\![g \vdash x \simeq xe;$
$\quad g \vdash xs \simeq_L xse]\!]$
$\quad \Longrightarrow g \vdash x\#xs \simeq_L xe\#xse$

code-pred $(modes: i \Rightarrow i \Rightarrow o \Rightarrow bool\ as\ exprListE)\ replist\ \langle proof \rangle$

definition $wf\text{-}term\text{-}graph :: MapState \Rightarrow Params \Rightarrow IRGraph \Rightarrow ID \Rightarrow bool$ **where**
$wf\text{-}term\text{-}graph\ m\ p\ g\ n = (\exists\ e.\ (g \vdash n \simeq e) \wedge (\exists\ v.\ ([m, p] \vdash e \mapsto v)))$

values $\{t.\ eg2\text{-}sq \vdash 4 \simeq t\}$

6.2 Data-flow Tree to Subgraph

fun $unary\text{-}node :: IRUnaryOp \Rightarrow ID \Rightarrow IRNode$ **where**
$unary\text{-}node\ UnaryAbs\ v = AbsNode\ v \mid$
$unary\text{-}node\ UnaryNot\ v = NotNode\ v \mid$
$unary\text{-}node\ UnaryNeg\ v = NegateNode\ v \mid$
$unary\text{-}node\ UnaryLogicNegation\ v = LogicNegationNode\ v \mid$
$unary\text{-}node\ (UnaryNarrow\ ib\ rb)\ v = NarrowNode\ ib\ rb\ v \mid$
$unary\text{-}node\ (UnarySignExtend\ ib\ rb)\ v = SignExtendNode\ ib\ rb\ v \mid$
$unary\text{-}node\ (UnaryZeroExtend\ ib\ rb)\ v = ZeroExtendNode\ ib\ rb\ v$

fun $bin\text{-}node :: IRBinaryOp \Rightarrow ID \Rightarrow ID \Rightarrow IRNode$ **where**
$bin\text{-}node\ BinAdd\ x\ y = AddNode\ x\ y \mid$
$bin\text{-}node\ BinMul\ x\ y = MulNode\ x\ y \mid$
$bin\text{-}node\ BinSub\ x\ y = SubNode\ x\ y \mid$
$bin\text{-}node\ BinAnd\ x\ y = AndNode\ x\ y \mid$
$bin\text{-}node\ BinOr\ \ x\ y = OrNode\ x\ y \mid$

bin-node BinXor x y = XorNode x y |
bin-node BinShortCircuitOr x y = ShortCircuitOrNode x y |
bin-node BinLeftShift x y = LeftShiftNode x y |
bin-node BinRightShift x y = RightShiftNode x y |
bin-node BinURightShift x y = UnsignedRightShiftNode x y |
bin-node BinIntegerEquals x y = IntegerEqualsNode x y |
bin-node BinIntegerLessThan x y = IntegerLessThanNode x y |
bin-node BinIntegerBelow x y = IntegerBelowNode x y

inductive *fresh-id* :: *IRGraph* ⇒ *ID* ⇒ *bool* **where**
n ∉ *ids g* ⟹ *fresh-id g n*

code-pred *fresh-id* ⟨*proof*⟩

fun *get-fresh-id* :: *IRGraph* ⇒ *ID* **where**

get-fresh-id g = *last*(*sorted-list-of-set*(*ids g*)) + *1*

export-code *get-fresh-id*

value *get-fresh-id eg2-sq*
value *get-fresh-id* (*add-node 6* (*ParameterNode 2*, *default-stamp*) *eg2-sq*)

inductive
unrep :: *IRGraph* ⇒ *IRExpr* ⇒ (*IRGraph* × *ID*) ⇒ *bool* (- ⊕ - ⇝ - *55*)
where

ConstantNodeSame:
⟦*find-node-and-stamp g* (*ConstantNode c*, *constantAsStamp c*) = *Some n*⟧
⟹ *g* ⊕ (*ConstantExpr c*) ⇝ (*g*, *n*) |

ConstantNodeNew:
⟦*find-node-and-stamp g* (*ConstantNode c*, *constantAsStamp c*) = *None*;
n = *get-fresh-id g*;
g′ = *add-node n* (*ConstantNode c*, *constantAsStamp c*) *g* ⟧
⟹ *g* ⊕ (*ConstantExpr c*) ⇝ (*g′*, *n*) |

ParameterNodeSame:
⟦*find-node-and-stamp g* (*ParameterNode i*, *s*) = *Some n*⟧
⟹ *g* ⊕ (*ParameterExpr i s*) ⇝ (*g*, *n*) |

ParameterNodeNew:
⟦*find-node-and-stamp g* (*ParameterNode i*, *s*) = *None*;
n = *get-fresh-id g*;
g′ = *add-node n* (*ParameterNode i*, *s*) *g*⟧

$\Longrightarrow g \oplus (\mathit{ParameterExpr}\ i\ s) \rightsquigarrow (g', n)\ |$

ConditionalNodeSame:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g4\ (\mathit{ConditionalNode}\ c\ t\ f,\ s') = \mathit{Some}\ n;$
$g \oplus ce \rightsquigarrow (g2, c);$
$g2 \oplus te \rightsquigarrow (g3, t);$
$g3 \oplus fe \rightsquigarrow (g4, f);$
$s' = \mathit{meet}\ (\mathit{stamp}\ g4\ t)\ (\mathit{stamp}\ g4\ f)]\!]$
$\Longrightarrow g \oplus (\mathit{ConditionalExpr}\ ce\ te\ fe) \rightsquigarrow (g4, n)\ |$

ConditionalNodeNew:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g4\ (\mathit{ConditionalNode}\ c\ t\ f,\ s') = \mathit{None};$
$g \oplus ce \rightsquigarrow (g2, c);$
$g2 \oplus te \rightsquigarrow (g3, t);$
$g3 \oplus fe \rightsquigarrow (g4, f);$
$s' = \mathit{meet}\ (\mathit{stamp}\ g4\ t)\ (\mathit{stamp}\ g4\ f);$
$n = \mathit{get\text{-}fresh\text{-}id}\ g4;$
$g' = \mathit{add\text{-}node}\ n\ (\mathit{ConditionalNode}\ c\ t\ f,\ s')\ g4]\!]$
$\Longrightarrow g \oplus (\mathit{ConditionalExpr}\ ce\ te\ fe) \rightsquigarrow (g', n)\ |$

UnaryNodeSame:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g2\ (\mathit{unary\text{-}node}\ op\ x,\ s') = \mathit{Some}\ n;$
$g \oplus xe \rightsquigarrow (g2, x);$
$s' = \mathit{stamp\text{-}unary}\ op\ (\mathit{stamp}\ g2\ x)]\!]$
$\Longrightarrow g \oplus (\mathit{UnaryExpr}\ op\ xe) \rightsquigarrow (g2, n)\ |$

UnaryNodeNew:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g2\ (\mathit{unary\text{-}node}\ op\ x,\ s') = \mathit{None};$
$g \oplus xe \rightsquigarrow (g2, x);$
$s' = \mathit{stamp\text{-}unary}\ op\ (\mathit{stamp}\ g2\ x);$
$n = \mathit{get\text{-}fresh\text{-}id}\ g2;$
$g' = \mathit{add\text{-}node}\ n\ (\mathit{unary\text{-}node}\ op\ x,\ s')\ g2]\!]$
$\Longrightarrow g \oplus (\mathit{UnaryExpr}\ op\ xe) \rightsquigarrow (g', n)\ |$

BinaryNodeSame:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g3\ (\mathit{bin\text{-}node}\ op\ x\ y,\ s') = \mathit{Some}\ n;$
$g \oplus xe \rightsquigarrow (g2, x);$
$g2 \oplus ye \rightsquigarrow (g3, y);$
$s' = \mathit{stamp\text{-}binary}\ op\ (\mathit{stamp}\ g3\ x)\ (\mathit{stamp}\ g3\ y)]\!]$
$\Longrightarrow g \oplus (\mathit{BinaryExpr}\ op\ xe\ ye) \rightsquigarrow (g3, n)\ |$

BinaryNodeNew:
$[\![\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g3\ (\mathit{bin\text{-}node}\ op\ x\ y,\ s') = \mathit{None};$
$g \oplus xe \rightsquigarrow (g2, x);$
$g2 \oplus ye \rightsquigarrow (g3, y);$
$s' = \mathit{stamp\text{-}binary}\ op\ (\mathit{stamp}\ g3\ x)\ (\mathit{stamp}\ g3\ y);$
$n = \mathit{get\text{-}fresh\text{-}id}\ g3;$
$g' = \mathit{add\text{-}node}\ n\ (\mathit{bin\text{-}node}\ op\ x\ y,\ s')\ g3]\!]$
$\Longrightarrow g \oplus (\mathit{BinaryExpr}\ op\ xe\ ye) \rightsquigarrow (g', n)\ |$

$AllLeafNodes$:
$[\![stamp\ g\ n = s;$
$is\text{-}preevaluated\ (kind\ g\ n)]\!]$
$\Longrightarrow g \oplus (LeafExpr\ n\ s) \rightsquigarrow (g,\ n)$

code-pred $(modes\colon i \Rightarrow i \Rightarrow o \Rightarrow bool\ as\ unrepE)$
$unrep\ \langle proof \rangle$

unrepRules

$$\frac{\textit{find-node-and-stamp}\ g\ (\textit{ConstantNode}\ c,\ \textit{constantAsStamp}\ c) = \textit{Some}\ n}{g \oplus \textit{ConstantExpr}\ c \leadsto (g,\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g\ (\textit{ConstantNode}\ c,\ \textit{constantAsStamp}\ c) = \textit{None} \\ n = \textit{get-fresh-id}\ g \\ g' = \textit{add-node}\ n\ (\textit{ConstantNode}\ c,\ \textit{constantAsStamp}\ c)\ g\end{array}}{g \oplus \textit{ConstantExpr}\ c \leadsto (g',\ n)}$$

$$\frac{\textit{find-node-and-stamp}\ g\ (\textit{ParameterNode}\ i,\ s) = \textit{Some}\ n}{g \oplus \textit{ParameterExpr}\ i\ s \leadsto (g,\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g\ (\textit{ParameterNode}\ i,\ s) = \textit{None} \\ n = \textit{get-fresh-id}\ g \qquad g' = \textit{add-node}\ n\ (\textit{ParameterNode}\ i,\ s)\ g\end{array}}{g \oplus \textit{ParameterExpr}\ i\ s \leadsto (g',\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g4\ (\textit{ConditionalNode}\ c\ t\ f,\ s') = \textit{Some}\ n \\ g \oplus ce \leadsto (g2,\ c) \qquad g2 \oplus te \leadsto (g3,\ t) \\ g3 \oplus fe \leadsto (g4,\ f) \qquad s' = \textit{meet}\ (\textit{stamp}\ g4\ t)\ (\textit{stamp}\ g4\ f)\end{array}}{g \oplus \textit{ConditionalExpr}\ ce\ te\ fe \leadsto (g4,\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g4\ (\textit{ConditionalNode}\ c\ t\ f,\ s') = \textit{None} \\ g \oplus ce \leadsto (g2,\ c) \qquad g2 \oplus te \leadsto (g3,\ t) \\ g3 \oplus fe \leadsto (g4,\ f) \qquad s' = \textit{meet}\ (\textit{stamp}\ g4\ t)\ (\textit{stamp}\ g4\ f) \\ n = \textit{get-fresh-id}\ g4 \qquad g' = \textit{add-node}\ n\ (\textit{ConditionalNode}\ c\ t\ f,\ s')\ g4\end{array}}{g \oplus \textit{ConditionalExpr}\ ce\ te\ fe \leadsto (g',\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g3\ (\textit{bin-node}\ op\ x\ y,\ s') = \textit{Some}\ n \\ g \oplus xe \leadsto (g2,\ x) \\ g2 \oplus ye \leadsto (g3,\ y) \qquad s' = \textit{stamp-binary}\ op\ (\textit{stamp}\ g3\ x)\ (\textit{stamp}\ g3\ y)\end{array}}{g \oplus \textit{BinaryExpr}\ op\ xe\ ye \leadsto (g3,\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g3\ (\textit{bin-node}\ op\ x\ y,\ s') = \textit{None} \\ g \oplus xe \leadsto (g2,\ x) \\ g2 \oplus ye \leadsto (g3,\ y) \qquad s' = \textit{stamp-binary}\ op\ (\textit{stamp}\ g3\ x)\ (\textit{stamp}\ g3\ y) \\ n = \textit{get-fresh-id}\ g3 \qquad g' = \textit{add-node}\ n\ (\textit{bin-node}\ op\ x\ y,\ s')\ g3\end{array}}{g \oplus \textit{BinaryExpr}\ op\ xe\ ye \leadsto (g',\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g2\ (\textit{unary-node}\ op\ x,\ s') = \textit{Some}\ n \\ g \oplus xe \leadsto (g2,\ x) \qquad s' = \textit{stamp-unary}\ op\ (\textit{stamp}\ g2\ x)\end{array}}{g \oplus \textit{UnaryExpr}\ op\ xe \leadsto (g2,\ n)}$$

$$\frac{\begin{array}{c}\textit{find-node-and-stamp}\ g2\ (\textit{unary-node}\ op\ x,\ s') = \textit{None} \\ g \oplus xe \leadsto (g2,\ x) \qquad s' = \textit{stamp-unary}\ op\ (\textit{stamp}\ g2\ x) \\ n = \textit{get-fresh-id}\ g2 \qquad g' = \textit{add-node}\ n\ (\textit{unary-node}\ op\ x,\ s')\ g2\end{array}}{g \oplus \textit{UnaryExpr}\ op\ xe \leadsto (g',\ n)}$$

$$\frac{\textit{stamp}\ g\ n = s \qquad \textit{is-preevaluated}\ (\textit{kind}\ g\ n)}{g \oplus \textit{LeafExpr}\ n\ s \leadsto (g,\ n)}$$

values $\{(n,\ g)\ .\ (eg2\text{-}sq \oplus sq\text{-}param0 \rightsquigarrow (g,\ n))\}$

6.3 Lift Data-flow Tree Semantics

definition *encodeeval* :: *IRGraph* ⇒ *MapState* ⇒ *Params* ⇒ *ID* ⇒ *Value* ⇒ *bool*
([-,-,-] ⊢ - ↦ - *50*)
where
encodeeval g m p n v = (∃ *e*. (*g* ⊢ *n* ≃ *e*) ∧ ([*m*,*p*] ⊢ *e* ↦ *v*))

6.4 Graph Refinement

definition *graph-represents-expression* :: *IRGraph* ⇒ *ID* ⇒ *IRExpr* ⇒ *bool*
(- ⊢ - ⊴ - *50*)
where
(*g* ⊢ *n* ⊴ *e*) = (∃ *e′* . (*g* ⊢ *n* ≃ *e′*) ∧ (*e′* ≤ *e*))

definition *graph-refinement* :: *IRGraph* ⇒ *IRGraph* ⇒ *bool* **where**
graph-refinement g_1 g_2 =
((*ids* g_1 ⊆ *ids* g_2) ∧
(∀ *n* . *n* ∈ *ids* g_1 ⟶ (∀ *e*. (g_1 ⊢ *n* ≃ *e*) ⟶ (g_2 ⊢ *n* ⊴ *e*))))

lemma *graph-refinement*:
graph-refinement g1 g2 ⟹ (∀ *n m p v*. *n* ∈ *ids g1* ⟶ ([*g1*, *m*, *p*] ⊢ *n* ↦ *v*) ⟶ ([*g2*, *m*, *p*] ⊢ *n* ↦ *v*))
⟨*proof*⟩

6.5 Maximal Sharing

definition *maximal-sharing*:
maximal-sharing g = (∀ n_1 n_2 . n_1 ∈ *true-ids g* ∧ n_2 ∈ *true-ids g* ⟶
(∀ *e*. (*g* ⊢ n_1 ≃ *e*) ∧ (*g* ⊢ n_2 ≃ *e*) ∧ (*stamp g* n_1 = *stamp g* n_2) ⟶ n_1 = n_2))

end

6.6 Formedness Properties

theory *Form*
imports
Semantics.TreeToGraph
begin

definition *wf-start* **where**
wf-start g = (*0* ∈ *ids g* ∧
is-StartNode (*kind g 0*))

definition *wf-closed* **where**
wf-closed g =
(∀ *n* ∈ *ids g* .
inputs g n ⊆ *ids g* ∧

succ g n $\subseteq$ *ids g* $\wedge$
kind g n $\neq$ *NoNode*)

definition *wf-phis* **where**
wf-phis g =
($\forall$ *n* $\in$ *ids g*.
is-PhiNode (*kind g n*) $\longrightarrow$
length (*ir-values* (*kind g n*))
= *length* (*ir-ends*
(*kind g* (*ir-merge* (*kind g n*)))))

definition *wf-ends* **where**
wf-ends g =
($\forall$ *n* $\in$ *ids g* .
is-AbstractEndNode (*kind g n*) $\longrightarrow$
card (*usages g n*) > *0*)

fun *wf-graph* :: *IRGraph* $\Rightarrow$ *bool* **where**
wf-graph g = (*wf-start g* $\wedge$ *wf-closed g* $\wedge$ *wf-phis g* $\wedge$ *wf-ends g*)

lemmas *wf-folds* =
wf-graph.simps
wf-start-def
wf-closed-def
wf-phis-def
wf-ends-def

fun *wf-stamps* :: *IRGraph* $\Rightarrow$ *bool* **where**
wf-stamps g = ($\forall$ *n* $\in$ *ids g* .
($\forall$ *v m p e* . (*g* $\vdash$ *n* $\simeq$ *e*) $\wedge$ ([*m*, *p*] $\vdash$ *e* $\mapsto$ *v*) $\longrightarrow$ *valid-value v* (*stamp-expr e*)))

fun *wf-stamp* :: *IRGraph* $\Rightarrow$ (*ID* $\Rightarrow$ *Stamp*) $\Rightarrow$ *bool* **where**
wf-stamp g s = ($\forall$ *n* $\in$ *ids g* .
($\forall$ *v m p e* . (*g* $\vdash$ *n* $\simeq$ *e*) $\wedge$ ([*m*, *p*] $\vdash$ *e* $\mapsto$ *v*) $\longrightarrow$ *valid-value v* (*s n*)))

lemma *wf-empty*: *wf-graph start-end-graph*
⟨*proof*⟩

lemma *wf-eg2-sq*: *wf-graph eg2-sq*
⟨*proof*⟩

fun *wf-logic-node-inputs* :: *IRGraph* $\Rightarrow$ *ID* $\Rightarrow$ *bool* **where**
wf-logic-node-inputs g n =
($\forall$ *inp* $\in$ *set* (*inputs-of* (*kind g n*)) . ($\forall$ *v m p* . ([*g*, *m*, *p*] $\vdash$ *inp* $\mapsto$ *v*) $\longrightarrow$ *wf-bool v*))

fun *wf-values* :: *IRGraph* $\Rightarrow$ *bool* **where**
wf-values g = ($\forall$ *n* $\in$ *ids g* .
($\forall$ *v m p* . ([*g*, *m*, *p*] $\vdash$ *n* $\mapsto$ *v*) $\longrightarrow$

($is\text{-}LogicNode$ ($kind\ g\ n$) $\longrightarrow$
$wf\text{-}bool\ v \wedge wf\text{-}logic\text{-}node\text{-}inputs\ g\ n$)))

end

6.7 Dynamic Frames

This theory defines two operators, 'unchanged' and 'changeonly', that are useful for specifying which nodes in an IRGraph can change. The dynamic framing idea originates from 'Dynamic Frames' in software verification, started by Ioannis T. Kassios in "Dynamic frames: Support for framing, dependencies and sharing without restrictions", In FM 2006.

theory *IRGraphFrames*
imports
Form
begin

fun *unchanged* :: $ID\ set \Rightarrow IRGraph \Rightarrow IRGraph \Rightarrow bool$ **where**
$unchanged\ ns\ g1\ g2 = (\forall\ n\ .\ n \in ns \longrightarrow$
$(n \in ids\ g1 \wedge n \in ids\ g2 \wedge kind\ g1\ n = kind\ g2\ n \wedge stamp\ g1\ n = stamp\ g2\ n))$

fun *changeonly* :: $ID\ set \Rightarrow IRGraph \Rightarrow IRGraph \Rightarrow bool$ **where**
$changeonly\ ns\ g1\ g2 = (\forall\ n\ .\ n \in ids\ g1 \wedge n \notin ns \longrightarrow$
$(n \in ids\ g1 \wedge n \in ids\ g2 \wedge kind\ g1\ n = kind\ g2\ n \wedge stamp\ g1\ n = stamp\ g2\ n))$

lemma *node-unchanged*:
assumes $unchanged\ ns\ g1\ g2$
assumes $nid \in ns$
shows $kind\ g1\ nid = kind\ g2\ nid$
⟨*proof*⟩

lemma *other-node-unchanged*:
assumes $changeonly\ ns\ g1\ g2$
assumes $nid \in ids\ g1$
assumes $nid \notin ns$
shows $kind\ g1\ nid = kind\ g2\ nid$
⟨*proof*⟩

Some notation for input nodes used

inductive *eval-uses*:: $IRGraph \Rightarrow ID \Rightarrow ID \Rightarrow bool$
for g **where**

use0: $nid \in ids\ g$
$\Longrightarrow eval\text{-}uses\ g\ nid\ nid$ |

use-inp: $nid' \in inputs\ g\ n$
$\Longrightarrow eval\text{-}uses\ g\ nid\ nid'$ |

use-trans: $[\![$*eval-uses g nid nid'*;
eval-uses g nid' nid''$]\!]$
$\Longrightarrow$ *eval-uses g nid nid''*

fun *eval-usages* :: *IRGraph* $\Rightarrow$ *ID* $\Rightarrow$ *ID set* **where**
eval-usages g nid = $\{$*n* $\in$ *ids g* . *eval-uses g nid n*$\}$

lemma *eval-usages-self*:
assumes *nid* $\in$ *ids g*
shows *nid* $\in$ *eval-usages g nid*
⟨*proof*⟩

lemma *not-in-g-inputs*:
assumes *nid* $\notin$ *ids g*
shows *inputs g nid* = $\{\}$
⟨*proof*⟩

lemma *child-member*:
assumes *n* = *kind g nid*
assumes *n* $\neq$ *NoNode*
assumes *List.member* (*inputs-of n*) *child*
shows *child* $\in$ *inputs g nid*
⟨*proof*⟩

lemma *child-member-in*:
assumes *nid* $\in$ *ids g*
assumes *List.member* (*inputs-of* (*kind g nid*)) *child*
shows *child* $\in$ *inputs g nid*
⟨*proof*⟩

lemma *inp-in-g*:
assumes *n* $\in$ *inputs g nid*
shows *nid* $\in$ *ids g*
⟨*proof*⟩

lemma *inp-in-g-wf*:
assumes *wf-graph g*
assumes *n* $\in$ *inputs g nid*
shows *n* $\in$ *ids g*
⟨*proof*⟩

lemma *kind-unchanged*:
assumes *nid* $\in$ *ids g1*
assumes *unchanged* (*eval-usages g1 nid*) *g1 g2*

shows *kind g1 nid* = *kind g2 nid*
⟨*proof*⟩

lemma *stamp-unchanged*:
assumes *nid* ∈ *ids g1*
assumes *unchanged* (*eval-usages g1 nid*) *g1 g2*
shows *stamp g1 nid* = *stamp g2 nid*
⟨*proof*⟩

lemma *child-unchanged*:
assumes *child* ∈ *inputs g1 nid*
assumes *unchanged* (*eval-usages g1 nid*) *g1 g2*
shows *unchanged* (*eval-usages g1 child*) *g1 g2*
⟨*proof*⟩

lemma *eval-usages*:
assumes *us* = *eval-usages g nid*
assumes *nid'* ∈ *ids g*
shows *eval-uses g nid nid'* ⟷ *nid'* ∈ *us* (**is** *?P* ⟷ *?Q*)
⟨*proof*⟩

lemma *inputs-are-uses*:
assumes *nid'* ∈ *inputs g nid*
shows *eval-uses g nid nid'*
⟨*proof*⟩

lemma *inputs-are-usages*:
assumes *nid'* ∈ *inputs g nid*
assumes *nid'* ∈ *ids g*
shows *nid'* ∈ *eval-usages g nid*
⟨*proof*⟩

lemma *inputs-of-are-usages*:
assumes *List.member* (*inputs-of* (*kind g nid*)) *nid'*
assumes *nid'* ∈ *ids g*
shows *nid'* ∈ *eval-usages g nid*
⟨*proof*⟩

lemma *usage-includes-inputs*:
assumes *us* = *eval-usages g nid*
assumes *ls* = *inputs g nid*
assumes *ls* ⊆ *ids g*
shows *ls* ⊆ *us*
⟨*proof*⟩

lemma *elim-inp-set*:
assumes *k* = *kind g nid*
assumes *k* ≠ *NoNode*

assumes *child* $\in$ *set* (*inputs-of k*)
shows *child* $\in$ *inputs g nid*
⟨*proof*⟩

lemma *encode-in-ids*:
assumes $g \vdash nid \simeq e$
shows *nid* $\in$ *ids g*
⟨*proof*⟩

lemma *eval-in-ids*:
assumes $[g, m, p] \vdash nid \mapsto v$
shows *nid* $\in$ *ids g*
⟨*proof*⟩

lemma *transitive-kind-same*:
assumes *unchanged* (*eval-usages g1 nid*) *g1 g2*
shows $\forall$ *nid'* $\in$ (*eval-usages g1 nid*) . *kind g1 nid'* = *kind g2 nid'*
⟨*proof*⟩

theorem *stay-same-encoding*:
assumes *nc*: *unchanged* (*eval-usages g1 nid*) *g1 g2*
assumes *g1*: $g1 \vdash nid \simeq e$
assumes *wf*: *wf-graph g1*
shows $g2 \vdash nid \simeq e$
⟨*proof*⟩

theorem *stay-same*:
assumes *nc*: *unchanged* (*eval-usages g1 nid*) *g1 g2*
assumes *g1*: $[g1, m, p] \vdash nid \mapsto v1$
assumes *wf*: *wf-graph g1*
shows $[g2, m, p] \vdash nid \mapsto v1$
⟨*proof*⟩

lemma *add-changed*:
assumes *gup* = *add-node new k g*
shows *changeonly* {*new*} *g gup*
⟨*proof*⟩

lemma *disjoint-change*:
assumes *changeonly change g gup*
assumes *nochange* = *ids g* − *change*
shows *unchanged nochange g gup*
⟨*proof*⟩

lemma *add-node-unchanged*:
assumes *new* $\notin$ *ids g*

assumes $nid \in ids\ g$
assumes $gup = add\text{-}node\ new\ k\ g$
assumes $wf\text{-}graph\ g$
shows $unchanged\ (eval\text{-}usages\ g\ nid)\ g\ gup$
$\langle proof \rangle$

lemma $eval\text{-}uses\text{-}imp$:
$((nid' \in ids\ g \wedge nid = nid')$
$\vee\ nid' \in inputs\ g\ nid$
$\vee\ (\exists\, nid''\ .\ eval\text{-}uses\ g\ nid\ nid'' \wedge eval\text{-}uses\ g\ nid''\ nid'))$
$\longleftrightarrow eval\text{-}uses\ g\ nid\ nid'$
$\langle proof \rangle$

lemma $wf\text{-}use\text{-}ids$:
assumes $wf\text{-}graph\ g$
assumes $nid \in ids\ g$
assumes $eval\text{-}uses\ g\ nid\ nid'$
shows $nid' \in ids\ g$
$\langle proof \rangle$

lemma $no\text{-}external\text{-}use$:
assumes $wf\text{-}graph\ g$
assumes $nid' \notin ids\ g$
assumes $nid \in ids\ g$
shows $\neg(eval\text{-}uses\ g\ nid\ nid')$
$\langle proof \rangle$

end

6.8 Tree to Graph Theorems

theory *TreeToGraphThms*
imports
IRTreeEvalThms
IRGraphFrames
HOL−Eisbach.Eisbach
HOL−Eisbach.Eisbach-Tools
begin

6.8.1 Extraction and Evaluation of Expression Trees is Deterministic.

First, we prove some extra rules that relate each type of IRNode to the corresponding IRExpr type that 'rep' will produce. These are very helpful for proving that 'rep' is deterministic.

named-theorems *rep*

lemma *rep-constant* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$

$kind\ g\ n = ConstantNode\ c \Longrightarrow$
$e = ConstantExpr\ c$
⟨*proof*⟩

lemma *rep-parameter* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = ParameterNode\ i \Longrightarrow$
$(\exists s.\ e = ParameterExpr\ i\ s)$
⟨*proof*⟩

lemma *rep-conditional* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = ConditionalNode\ c\ t\ f \Longrightarrow$
$(\exists\ ce\ te\ fe.\ e = ConditionalExpr\ ce\ te\ fe)$
⟨*proof*⟩

lemma *rep-abs* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = AbsNode\ x \Longrightarrow$
$(\exists xe.\ e = UnaryExpr\ UnaryAbs\ xe)$
⟨*proof*⟩

lemma *rep-not* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = NotNode\ x \Longrightarrow$
$(\exists xe.\ e = UnaryExpr\ UnaryNot\ xe)$
⟨*proof*⟩

lemma *rep-negate* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = NegateNode\ x \Longrightarrow$
$(\exists xe.\ e = UnaryExpr\ UnaryNeg\ xe)$
⟨*proof*⟩

lemma *rep-logicnegation* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = LogicNegationNode\ x \Longrightarrow$
$(\exists xe.\ e = UnaryExpr\ UnaryLogicNegation\ xe)$
⟨*proof*⟩

lemma *rep-add* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = AddNode\ x\ y \Longrightarrow$
$(\exists xe\ ye.\ e = BinaryExpr\ BinAdd\ xe\ ye)$
⟨*proof*⟩

lemma *rep-sub* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = SubNode\ x\ y \Longrightarrow$

$(\exists\, xe\ ye.\ e = BinaryExpr\ BinSub\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-mul* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = MulNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinMul\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-and* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = AndNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinAnd\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-or* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = OrNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinOr\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-xor* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = XorNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinXor\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-short-circuit-or* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = ShortCircuitOrNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinShortCircuitOr\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-left-shift* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = LeftShiftNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinLeftShift\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-right-shift* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = RightShiftNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinRightShift\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-unsigned-right-shift* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = UnsignedRightShiftNode\ x\ y \Longrightarrow$
$(\exists\, xe\ ye.\ e = BinaryExpr\ BinURightShift\ xe\ ye)$

$\langle proof \rangle$

lemma *rep-integer-below* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = IntegerBelowNode\ x\ y \Longrightarrow$
$(\exists xe\ ye.\ e = BinaryExpr\ BinIntegerBelow\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-integer-equals* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = IntegerEqualsNode\ x\ y \Longrightarrow$
$(\exists xe\ ye.\ e = BinaryExpr\ BinIntegerEquals\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-integer-less-than* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = IntegerLessThanNode\ x\ y \Longrightarrow$
$(\exists xe\ ye.\ e = BinaryExpr\ BinIntegerLessThan\ xe\ ye)$
$\langle proof \rangle$

lemma *rep-narrow* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = NarrowNode\ ib\ rb\ x \Longrightarrow$
$(\exists x.\ e = UnaryExpr\ (UnaryNarrow\ ib\ rb)\ x)$
$\langle proof \rangle$

lemma *rep-sign-extend* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = SignExtendNode\ ib\ rb\ x \Longrightarrow$
$(\exists x.\ e = UnaryExpr\ (UnarySignExtend\ ib\ rb)\ x)$
$\langle proof \rangle$

lemma *rep-zero-extend* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = ZeroExtendNode\ ib\ rb\ x \Longrightarrow$
$(\exists x.\ e = UnaryExpr\ (UnaryZeroExtend\ ib\ rb)\ x)$
$\langle proof \rangle$

lemma *rep-load-field* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$is\text{-}preevaluated\ (kind\ g\ n) \Longrightarrow$
$(\exists s.\ e = LeafExpr\ n\ s)$
$\langle proof \rangle$

lemma *rep-ref* [*rep*]:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n = RefNode\ n' \Longrightarrow$
$g \vdash n' \simeq e$
$\langle proof \rangle$

method *solve-det* **uses** *node* =
  (*match node* **in** *kind* - - = *node* - **for** *node* ⇒
    ‹*match rep in r*: - $\Longrightarrow$ - = *node* - $\Longrightarrow$ - ⇒
      ‹*match IRNode.inject in i*: (*node* - = *node* -) = - ⇒
        ‹*match RepE in e*: - $\Longrightarrow$ ($\bigwedge$*x*. - = *node x* $\Longrightarrow$ -) $\Longrightarrow$ - ⇒
          ‹*match IRNode.distinct in d*: *node* - ≠ *RefNode* - ⇒
            ‹*metis i e r d*›››› |
    match node **in** *kind* - - = *node* - - **for** *node* ⇒
    ‹*match rep in r*: - $\Longrightarrow$ - = *node* - - $\Longrightarrow$ - ⇒
      ‹*match IRNode.inject in i*: (*node* - - = *node* - -) = - ⇒
        ‹*match RepE in e*: - $\Longrightarrow$ ($\bigwedge$*x y*. - = *node x y* $\Longrightarrow$ -) $\Longrightarrow$ - ⇒
          ‹*match IRNode.distinct in d*: *node* - - ≠ *RefNode* - ⇒
            ‹*metis i e r d*›››› |
    match node **in** *kind* - - = *node* - - - **for** *node* ⇒
    ‹*match rep in r*: - $\Longrightarrow$ - = *node* - - - $\Longrightarrow$ - ⇒
      ‹*match IRNode.inject in i*: (*node* - - - = *node* - - -) = - ⇒
        ‹*match RepE in e*: - $\Longrightarrow$ ($\bigwedge$*x y z*. - = *node x y z* $\Longrightarrow$ -) $\Longrightarrow$ - ⇒
          ‹*match IRNode.distinct in d*: *node* - - - ≠ *RefNode* - ⇒
            ‹*metis i e r d*›››› |
  match node **in** *kind* - - = *node* - - - **for** *node* ⇒
    ‹*match rep in r*: - $\Longrightarrow$ - = *node* - - - $\Longrightarrow$ - ⇒
      ‹*match IRNode.inject in i*: (*node* - - - = *node* - - -) = - ⇒
        ‹*match RepE in e*: - $\Longrightarrow$ ($\bigwedge$*x*. - = *node* - - *x* $\Longrightarrow$ -) $\Longrightarrow$ - ⇒
          ‹*match IRNode.distinct in d*: *node* - - - ≠ *RefNode* - ⇒
            ‹*metis i e r d*›››››)

Now we can prove that 'rep' and 'eval', and their list versions, are deterministic.

lemma *repDet*:
  shows $(g \vdash n \simeq e_1) \Longrightarrow (g \vdash n \simeq e_2) \Longrightarrow e_1 = e_2$
⟨*proof*⟩

lemma *repAllDet*:
  $g \vdash xs \simeq_L e1 \Longrightarrow$
    $g \vdash xs \simeq_L e2 \Longrightarrow$
    $e1 = e2$
⟨*proof*⟩

lemma *encodeEvalDet*:
  $[g,m,p] \vdash e \mapsto v1 \Longrightarrow$
    $[g,m,p] \vdash e \mapsto v2 \Longrightarrow$
    $v1 = v2$
  ⟨*proof*⟩

lemma *graphDet*: $([g,m,p] \vdash n \mapsto v_1) \land ([g,m,p] \vdash n \mapsto v_2) \Longrightarrow v_1 = v_2$
  ⟨*proof*⟩

6.8.2 Monotonicity of Graph Refinement

Lift refinement monotonicity to graph level. Hopefully these shouldn't really be required.

lemma *mono-abs*:
  assumes $kind\ g1\ n = AbsNode\ x \wedge kind\ g2\ n = AbsNode\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$
  assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
  shows $e1 \geq e2$
  ⟨*proof*⟩

lemma *mono-not*:
  assumes $kind\ g1\ n = NotNode\ x \wedge kind\ g2\ n = NotNode\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$
  assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
  shows $e1 \geq e2$
  ⟨*proof*⟩

lemma *mono-negate*:
  assumes $kind\ g1\ n = NegateNode\ x \wedge kind\ g2\ n = NegateNode\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$
  assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
  shows $e1 \geq e2$
  ⟨*proof*⟩

lemma *mono-logic-negation*:
  assumes $kind\ g1\ n = LogicNegationNode\ x \wedge kind\ g2\ n = LogicNegationNode\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$
  assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
  shows $e1 \geq e2$
  ⟨*proof*⟩

lemma *mono-narrow*:
  assumes $kind\ g1\ n = NarrowNode\ ib\ rb\ x \wedge kind\ g2\ n = NarrowNode\ ib\ rb\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$
  assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
  shows $e1 \geq e2$
  ⟨*proof*⟩

lemma *mono-sign-extend*:
  assumes $kind\ g1\ n = SignExtendNode\ ib\ rb\ x \wedge kind\ g2\ n = SignExtendNode\ ib\ rb\ x$
  assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
  assumes $xe1 \geq xe2$

assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
shows $e1 \geq e2$
$\langle proof \rangle$

lemma *mono-zero-extend*:
assumes $kind\ g1\ n = ZeroExtendNode\ ib\ rb\ x \wedge kind\ g2\ n = ZeroExtendNode\ ib\ rb\ x$
assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
assumes $xe1 \geq xe2$
assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
shows $e1 \geq e2$
$\langle proof \rangle$

lemma *mono-conditional-graph*:
assumes $kind\ g1\ n = ConditionalNode\ c\ t\ f \wedge kind\ g2\ n = ConditionalNode\ c\ t\ f$
assumes $(g1 \vdash c \simeq ce1) \wedge (g2 \vdash c \simeq ce2)$
assumes $(g1 \vdash t \simeq te1) \wedge (g2 \vdash t \simeq te2)$
assumes $(g1 \vdash f \simeq fe1) \wedge (g2 \vdash f \simeq fe2)$
assumes $ce1 \geq ce2 \wedge te1 \geq te2 \wedge fe1 \geq fe2$
assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
shows $e1 \geq e2$
$\langle proof \rangle$

lemma *mono-add*:
assumes $kind\ g1\ n = AddNode\ x\ y \wedge kind\ g2\ n = AddNode\ x\ y$
assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
assumes $(g1 \vdash y \simeq ye1) \wedge (g2 \vdash y \simeq ye2)$
assumes $xe1 \geq xe2 \wedge ye1 \geq ye2$
assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
shows $e1 \geq e2$
$\langle proof \rangle$

lemma *mono-mul*:
assumes $kind\ g1\ n = MulNode\ x\ y \wedge kind\ g2\ n = MulNode\ x\ y$
assumes $(g1 \vdash x \simeq xe1) \wedge (g2 \vdash x \simeq xe2)$
assumes $(g1 \vdash y \simeq ye1) \wedge (g2 \vdash y \simeq ye2)$
assumes $xe1 \geq xe2 \wedge ye1 \geq ye2$
assumes $(g1 \vdash n \simeq e1) \wedge (g2 \vdash n \simeq e2)$
shows $e1 \geq e2$
$\langle proof \rangle$

lemma *term-graph-evaluation*:
$(g \vdash n \trianglelefteq e) \Longrightarrow (\forall\ m\ p\ v\ .\ ([m,p] \vdash e \mapsto v) \longrightarrow ([g,m,p] \vdash n \mapsto v))$
$\langle proof \rangle$

lemma *encodes-contains*:
$g \vdash n \simeq e \Longrightarrow$
$kind\ g\ n \neq NoNode$

$\langle proof \rangle$

lemma *no-encoding*:
assumes $n \notin ids\ g$
shows $\neg(g \vdash n \simeq e)$
$\langle proof \rangle$

lemma *not-excluded-keep-type*:
assumes $n \in ids\ g1$
assumes $n \notin excluded$
assumes $(excluded \trianglelefteq as\text{-}set\ g1) \subseteq as\text{-}set\ g2$
shows $kind\ g1\ n = kind\ g2\ n \wedge stamp\ g1\ n = stamp\ g2\ n$
$\langle proof \rangle$

method *metis-node-eq-unary* **for** $node :: {}'a \Rightarrow IRNode =$
$(match\ IRNode.inject$ **in** i: $(node\ \text{-} = node\ \text{-}) = \text{-} \Rightarrow$
$\langle metis\ i \rangle)$
method *metis-node-eq-binary* **for** $node :: {}'a \Rightarrow {}'a \Rightarrow IRNode =$
$(match\ IRNode.inject$ **in** i: $(node\ \text{-}\ \text{-} = node\ \text{-}\ \text{-}) = \text{-} \Rightarrow$
$\langle metis\ i \rangle)$
method *metis-node-eq-ternary* **for** $node :: {}'a \Rightarrow {}'a \Rightarrow {}'a \Rightarrow IRNode =$
$(match\ IRNode.inject$ **in** i: $(node\ \text{-}\ \text{-}\ \text{-} = node\ \text{-}\ \text{-}\ \text{-}) = \text{-} \Rightarrow$
$\langle metis\ i \rangle)$

6.8.3 Lift Data-flow Tree Refinement to Graph Refinement

theorem *graph-semantics-preservation*:
assumes a: $e1' \geq e2'$
assumes b: $(\{n'\} \trianglelefteq as\text{-}set\ g1) \subseteq as\text{-}set\ g2$
assumes c: $g1 \vdash n' \simeq e1'$
assumes d: $g2 \vdash n' \simeq e2'$
shows $graph\text{-}refinement\ g1\ g2$
$\langle proof \rangle$

lemma *graph-semantics-preservation-subscript*:
assumes a: $e_1' \geq e_2'$
assumes b: $(\{n\} \trianglelefteq as\text{-}set\ g_1) \subseteq as\text{-}set\ g_2$
assumes c: $g_1 \vdash n \simeq e_1'$
assumes d: $g_2 \vdash n \simeq e_2'$
shows $graph\text{-}refinement\ g_1\ g_2$
$\langle proof \rangle$

lemma *tree-to-graph-rewriting*:
$e_1 \geq e_2$
$\wedge (g_1 \vdash n \simeq e_1) \wedge maximal\text{-}sharing\ g_1$
$\wedge (\{n\} \trianglelefteq as\text{-}set\ g_1) \subseteq as\text{-}set\ g_2$
$\wedge (g_2 \vdash n \simeq e_2) \wedge maximal\text{-}sharing\ g_2$
$\Longrightarrow graph\text{-}refinement\ g_1\ g_2$

$\langle proof \rangle$

declare [[*simp-trace*]]
lemma *equal-refines*:
fixes *e1 e2* :: *IRExpr*
assumes $e1 = e2$
shows $e1 \geq e2$
$\langle proof \rangle$
declare [[*simp-trace=false*]]

lemma *eval-contains-id*[*simp*]: $g1 \vdash n \simeq e \Longrightarrow n \in ids\ g1$
$\langle proof \rangle$

lemma *subset-kind*[*simp*]: $as\text{-}set\ g1 \subseteq as\text{-}set\ g2 \Longrightarrow g1 \vdash n \simeq e \Longrightarrow kind\ g1\ n = kind\ g2\ n$
$\langle proof \rangle$

lemma *subset-stamp*[*simp*]: $as\text{-}set\ g1 \subseteq as\text{-}set\ g2 \Longrightarrow g1 \vdash n \simeq e \Longrightarrow stamp\ g1\ n = stamp\ g2\ n$
$\langle proof \rangle$

method *solve-subset-eval* **uses** *as-set eval* =
(*metis eval as-set subset-kind subset-stamp* |
metis eval as-set subset-kind)

lemma *subset-implies-evals*:
assumes $as\text{-}set\ g1 \subseteq as\text{-}set\ g2$
assumes $(g1 \vdash n \simeq e)$
shows $(g2 \vdash n \simeq e)$
$\langle proof \rangle$

lemma *subset-refines*:
assumes $as\text{-}set\ g1 \subseteq as\text{-}set\ g2$
shows *graph-refinement g1 g2*
$\langle proof \rangle$

lemma *graph-construction*:
$e_1 \geq e_2$
$\wedge\ as\text{-}set\ g_1 \subseteq as\text{-}set\ g_2$
$\wedge\ (g_2 \vdash n \simeq e_2)$
$\Longrightarrow (g_2 \vdash n \trianglelefteq e_1) \wedge graph\text{-}refinement\ g_1\ g_2$
$\langle proof \rangle$

6.8.4 Term Graph Reconstruction

lemma *find-exists-kind*:

assumes *find-node-and-stamp g* $(node, s) = Some\ nid$
shows *kind g nid* $=$ *node*
⟨*proof*⟩

lemma *find-exists-stamp*:
assumes *find-node-and-stamp g* $(node, s) = Some\ nid$
shows *stamp g nid* $= s$
⟨*proof*⟩

lemma *find-new-kind*:
assumes $g' = add\text{-}node\ nid\ (node, s)\ g$
assumes $node \neq NoNode$
shows $kind\ g'\ nid = node$
⟨*proof*⟩

lemma *find-new-stamp*:
assumes $g' = add\text{-}node\ nid\ (node, s)\ g$
assumes $node \neq NoNode$
shows $stamp\ g'\ nid = s$
⟨*proof*⟩

lemma *sorted-bottom*:
assumes *finite xs*
assumes $x \in xs$
shows $x \leq last(sorted\text{-}list\text{-}of\text{-}set(xs::nat\ set))$
⟨*proof*⟩

lemma *fresh*: $finite\ xs \Longrightarrow last(sorted\text{-}list\text{-}of\text{-}set(xs::nat\ set)) + 1 \notin xs$
⟨*proof*⟩

lemma *fresh-ids*:
assumes $n = get\text{-}fresh\text{-}id\ g$
shows $n \notin ids\ g$
⟨*proof*⟩

lemma *graph-unchanged-rep-unchanged*:
assumes $\forall\, n \in ids\ g.\ kind\ g\ n = kind\ g'\ n$
assumes $\forall\, n \in ids\ g.\ stamp\ g\ n = stamp\ g'\ n$
shows $(g \vdash n \simeq e) \longrightarrow (g' \vdash n \simeq e)$
⟨*proof*⟩

lemma *fresh-node-subset*:
assumes $n \notin ids\ g$
assumes $g' = add\text{-}node\ n\ (k, s)\ g$
shows $as\text{-}set\ g \subseteq as\text{-}set\ g'$
⟨*proof*⟩

lemma *unrep-subset*:
assumes $(g \oplus e \rightsquigarrow (g', n))$

shows $\mathit{as\text{-}set}\ g \subseteq \mathit{as\text{-}set}\ g'$
⟨*proof*⟩

lemma *fresh-node-preserves-other-nodes*:
assumes $n' = \mathit{get\text{-}fresh\text{-}id}\ g$
assumes $g' = \mathit{add\text{-}node}\ n'\ (k,\ s)\ g$
shows $\forall\ n \in \mathit{ids}\ g\ .\ (g \vdash n \simeq e) \longrightarrow (g' \vdash n \simeq e)$
⟨*proof*⟩

lemma *found-node-preserves-other-nodes*:
assumes $\mathit{find\text{-}node\text{-}and\text{-}stamp}\ g\ (k,\ s) = \mathit{Some}\ n$
shows $\forall\ n \in \mathit{ids}\ g.\ (g \vdash n \simeq e) \longleftrightarrow (g \vdash n \simeq e)$
⟨*proof*⟩

lemma *unrep-ids-subset*[*simp*]:
assumes $g \oplus e \leadsto (g',\ n)$
shows $\mathit{ids}\ g \subseteq \mathit{ids}\ g'$
⟨*proof*⟩

lemma *unrep-unchanged*:
assumes $g \oplus e \leadsto (g',\ n)$
shows $\forall\ n \in \mathit{ids}\ g\ .\ \forall\ e.\ (g \vdash n \simeq e) \longrightarrow (g' \vdash n \simeq e)$
⟨*proof*⟩

theorem *term-graph-reconstruction*:
$g \oplus e \leadsto (g',\ n) \Longrightarrow (g' \vdash n \simeq e) \wedge \mathit{as\text{-}set}\ g \subseteq \mathit{as\text{-}set}\ g'$
⟨*proof*⟩

lemma *ref-refinement*:
assumes $g \vdash n \simeq e_1$
assumes $\mathit{kind}\ g\ n' = \mathit{RefNode}\ n$
shows $g \vdash n' \trianglelefteq e_1$
⟨*proof*⟩

lemma *unrep-refines*:
assumes $g \oplus e \leadsto (g',\ n)$
shows $\mathit{graph\text{-}refinement}\ g\ g'$
⟨*proof*⟩

lemma *add-new-node-refines*:
assumes $n \notin \mathit{ids}\ g$
assumes $g' = \mathit{add\text{-}node}\ n\ (k,\ s)\ g$
shows $\mathit{graph\text{-}refinement}\ g\ g'$
⟨*proof*⟩

lemma *add-node-as-set*:
assumes $g' = \mathit{add\text{-}node}\ n\ (k,\ s)\ g$
shows $(\{n\} \trianglelefteq \mathit{as\text{-}set}\ g) \subseteq \mathit{as\text{-}set}\ g'$
⟨*proof*⟩

theorem *refined-insert*:
assumes $e_1 \geq e_2$
assumes $g_1 \oplus e_2 \rightsquigarrow (g_2, n')$
shows $(g_2 \vdash n' \trianglelefteq e_1) \wedge$ *graph-refinement* $g_1\ g_2$
⟨*proof*⟩

lemma *ids-finite*: *finite* (*ids* g)
⟨*proof*⟩

lemma *unwrap-sorted*: *set* (*sorted-list-of-set* (*ids* g)) = *ids* g
⟨*proof*⟩

lemma *find-none*:
assumes *find-node-and-stamp* g (k, s) = *None*
shows $\forall\ n \in$ *ids* g. *kind* $g\ n \neq k \vee$ *stamp* $g\ n \neq s$
⟨*proof*⟩

method *ref-represents* **uses** *node* =
(*metis IRNode.distinct(2755) RefNode dual-order.refl find-new-kind fresh-node-subset node subset-implies-evals*)

6.8.5 Data-flow Tree to Subgraph Preserves Maximal Sharing

lemma *same-kind-stamp-encodes-equal*:
assumes *kind* $g\ n$ = *kind* $g\ n'$
assumes *stamp* $g\ n$ = *stamp* $g\ n'$
assumes $\neg$(*is-preevaluated* (*kind* $g\ n$))
shows $\forall\ e. (g \vdash n \simeq e) \longrightarrow (g \vdash n' \simeq e)$
⟨*proof*⟩

lemma *new-node-not-present*:
assumes *find-node-and-stamp* g (*node*, s) = *None*
assumes n = *get-fresh-id* g
assumes g' = *add-node* n (*node*, s) g
shows $\forall\ n' \in$ *true-ids* g. $(\forall e. ((g \vdash n \simeq e) \wedge (g \vdash n' \simeq e)) \longrightarrow n = n')$

$\langle proof \rangle$

lemma *true-ids-def*:
true-ids $g = \{n \in ids\ g.\ \neg(is\text{-}RefNode\ (kind\ g\ n)) \wedge ((kind\ g\ n) \neq NoNode)\}$
$\langle proof \rangle$

lemma *add-node-some-node-def*:
assumes $k \neq NoNode$
assumes $g' = add\text{-}node\ nid\ (k,\ s)\ g$
shows $g' = Abs\text{-}IRGraph\ ((Rep\text{-}IRGraph\ g)(nid \mapsto (k,\ s)))$
$\langle proof \rangle$

lemma *ids-add-update-v1*:
assumes $g' = add\text{-}node\ nid\ (k,\ s)\ g$
assumes $k \neq NoNode$
shows $dom\ (Rep\text{-}IRGraph\ g') = dom\ (Rep\text{-}IRGraph\ g) \cup \{nid\}$
$\langle proof \rangle$

lemma *ids-add-update-v2*:
assumes $g' = add\text{-}node\ nid\ (k,\ s)\ g$
assumes $k \neq NoNode$
shows $nid \in ids\ g'$
$\langle proof \rangle$

lemma *add-node-ids-subset*:
assumes $n \in ids\ g$
assumes $g' = add\text{-}node\ n\ node\ g$
shows $ids\ g' = ids\ g \cup \{n\}$
$\langle proof \rangle$

lemma *convert-maximal*:
assumes $\forall\, n\ n'.\ n \in true\text{-}ids\ g \wedge n' \in true\text{-}ids\ g \longrightarrow (\forall\, e\ e'.\ (g \vdash n \simeq e) \wedge (g \vdash n' \simeq e') \longrightarrow e \neq e')$
shows *maximal-sharing g*
$\langle proof \rangle$

lemma *add-node-set-eq*:
assumes $k \neq NoNode$
assumes $n \notin ids\ g$
shows $as\text{-}set\ (add\text{-}node\ n\ (k,\ s)\ g) = as\text{-}set\ g \cup \{(n,\ (k,\ s))\}$
$\langle proof \rangle$

lemma *add-node-as-set-eq*:
assumes $g' = add\text{-}node\ n\ (k,\ s)\ g$
assumes $n \notin ids\ g$
shows $(\{n\} \trianglelefteq as\text{-}set\ g') = as\text{-}set\ g$
$\langle proof \rangle$

lemma *true-ids*:

$true\text{-}ids\ g = ids\ g - \{n \in ids\ g.\ is\text{-}RefNode\ (kind\ g\ n)\}$
$\langle proof \rangle$

lemma *as-set-ids*:
assumes $as\text{-}set\ g = as\text{-}set\ g'$
shows $ids\ g = ids\ g'$
$\langle proof \rangle$

lemma *ids-add-update*:
assumes $k \neq NoNode$
assumes $n \notin ids\ g$
assumes $g' = add\text{-}node\ n\ (k,\ s)\ g$
shows $ids\ g' = ids\ g \cup \{n\}$
$\langle proof \rangle$

lemma *true-ids-add-update*:
assumes $k \neq NoNode$
assumes $n \notin ids\ g$
assumes $g' = add\text{-}node\ n\ (k,\ s)\ g$
assumes $\neg(is\text{-}RefNode\ k)$
shows $true\text{-}ids\ g' = true\text{-}ids\ g \cup \{n\}$
$\langle proof \rangle$

lemma *new-def*:
assumes $(new \trianglelefteq as\text{-}set\ g') = as\text{-}set\ g$
shows $n \in ids\ g \longrightarrow n \notin new$
$\langle proof \rangle$

lemma *add-preserves-rep*:
assumes *unchanged*: $(new \trianglelefteq as\text{-}set\ g') = as\text{-}set\ g$
assumes *closed*: $wf\text{-}closed\ g$
assumes *existed*: $n \in ids\ g$
assumes $g' \vdash n \simeq e$
shows $g \vdash n \simeq e$
$\langle proof \rangle$

lemma *not-in-no-rep*:
$n \notin ids\ g \Longrightarrow \forall\, e.\ \neg(g \vdash n \simeq e)$
$\langle proof \rangle$

lemma *unary-inputs*:
assumes $kind\ g\ n = unary\text{-}node\ op\ x$
shows $inputs\ g\ n = \{x\}$
$\langle proof \rangle$

lemma *unary-succ*:
  assumes *kind g n* = *unary-node op x*
  shows *succ g n* = {}
  ⟨*proof*⟩

lemma *binary-inputs*:
  assumes *kind g n* = *bin-node op x y*
  shows *inputs g n* = {*x*, *y*}
  ⟨*proof*⟩

lemma *binary-succ*:
  assumes *kind g n* = *bin-node op x y*
  shows *succ g n* = {}
  ⟨*proof*⟩

lemma *unrep-contains*:
  assumes $g \oplus e \rightsquigarrow (g', n)$
  shows $n \in ids\ g'$
  ⟨*proof*⟩

lemma *unrep-preserves-contains*:
  assumes $n \in ids\ g$
  assumes $g \oplus e \rightsquigarrow (g', n')$
  shows $n \in ids\ g'$
  ⟨*proof*⟩

lemma *unrep-preserves-closure*:
  assumes *wf-closed g*
  assumes $g \oplus e \rightsquigarrow (g', n)$
  shows *wf-closed* g'
  ⟨*proof*⟩

inductive-cases *ConstUnrepE*: $g \oplus (ConstantExpr\ x) \rightsquigarrow (g', n)$

definition *constant-value* **where**
  constant-value = (*IntVal 32 0*)
definition *bad-graph* **where**
  bad-graph = *irgraph* [
    (*0*, *AbsNode 1*, *constantAsStamp constant-value*),
    (*1*, *RefNode 2*, *constantAsStamp constant-value*),
    (*2*, *ConstantNode constant-value*, *constantAsStamp constant-value*)
  ]

end

7 Control-flow Semantics

theory *IRStepObj*
  imports
    TreeToGraph
begin

7.1 Object Heap

The heap model we introduce maps field references to object instances to runtime values. We use the H[f][p] heap representation. See \cite{heap−reps−2011}. We also introduce the DynamicHeap type which allocates new object references sequentially storing the next free object reference as 'Free'.

heapdef

type-synonym $('a, 'b)$ *Heap* $=$ $'a \Rightarrow 'b \Rightarrow$ *Value*
type-synonym *Free* $=$ *nat*
type-synonym $('a, 'b)$ *DynamicHeap* $= ('a, 'b)$ *Heap* $\times$ *Free*

fun *h-load-field* :: $'a \Rightarrow 'b \Rightarrow ('a, 'b)$ *DynamicHeap* $\Rightarrow$ *Value* **where**
  h-load-field f r $(h, n) = h\ f\ r$

fun *h-store-field* :: $'a \Rightarrow 'b \Rightarrow$ *Value* $\Rightarrow ('a, 'b)$ *DynamicHeap* $\Rightarrow ('a, 'b)$ *DynamicHeap* **where**
  h-store-field f r v $(h, n) = (h(f := ((h\ f)(r := v))), n)$

fun *h-new-inst* :: $('a, 'b)$ *DynamicHeap* $\Rightarrow ('a, 'b)$ *DynamicHeap* $\times$ *Value* **where**
  h-new-inst $(h, n) = ((h,n+1), ($*ObjRef* $($*Some* $n)))$

type-synonym *FieldRefHeap* $= ($*string*, *objref*$)$ *DynamicHeap*

definition new-heap :: $('a, 'b)$ *DynamicHeap* **where**
  new-heap $= ((\lambda f.\ \lambda p.\ UndefVal), 0)$

7.2 Intraprocedural Semantics

fun *find-index* :: $'a \Rightarrow 'a$ *list* $\Rightarrow$ *nat* **where**
  find-index - $[] = 0$ |
  find-index v $(x \# xs) = ($*if* $(x=v)$ *then* 0 *else find-index v xs* $+ 1)$

fun *phi-list* :: *IRGraph* $\Rightarrow$ *ID* $\Rightarrow$ *ID list* **where**
  phi-list g n $=$
    $($*filter* $(\lambda x.($*is-PhiNode* $($*kind g x*$)))$
      $($*sorted-list-of-set* $($*usages g n*$)))$

fun *input-index* :: *IRGraph* $\Rightarrow$ *ID* $\Rightarrow$ *ID* $\Rightarrow$ *nat* **where**
  input-index g n n' $=$ *find-index n'* $($*inputs-of* $($*kind g n*$))$

fun *phi-inputs* :: *IRGraph* ⇒ *nat* ⇒ *ID list* ⇒ *ID list* **where**
phi-inputs g i nodes = (*map* (λ*n*. (*inputs-of* (*kind g n*))!(*i* + *1*)) *nodes*)

fun *set-phis* :: *ID list* ⇒ *Value list* ⇒ *MapState* ⇒ *MapState* **where**
set-phis [] [] *m* = *m* |
set-phis (*n* # *xs*) (*v* # *vs*) *m* = (*set-phis xs vs* (*m*(*n* := *v*))) |
set-phis [] (*v* # *vs*) *m* = *m* |
set-phis (*x* # *xs*) [] *m* = *m*

Intraprocedural semantics are given as a small-step semantics.

Within the context of a graph, the configuration triple, (ID, MethodState, Heap), is related to the subsequent configuration.

inductive *step* :: *IRGraph* ⇒ *Params* ⇒ (*ID* × *MapState* × *FieldRefHeap*) ⇒ (*ID* × *MapState* × *FieldRefHeap*) ⇒ *bool*
(-, - ⊢ - → - *55*) **for** *g p* **where**

SequentialNode:
⟦*is-sequential-node* (*kind g nid*);
nid' = (*successors-of* (*kind g nid*))!*0*⟧
⟹ *g*, *p* ⊢ (*nid*, *m*, *h*) → (*nid'*, *m*, *h*) |

IfNode:
⟦*kind g nid* = (*IfNode cond tb fb*);
g ⊢ *cond* ≃ *condE*;
[*m*, *p*] ⊢ *condE* ↦ *val*;
nid' = (*if val-to-bool val then tb else fb*)⟧
⟹ *g*, *p* ⊢ (*nid*, *m*, *h*) → (*nid'*, *m*, *h*) |

EndNodes:
⟦*is-AbstractEndNode* (*kind g nid*);
merge = *any-usage g nid*;
is-AbstractMergeNode (*kind g merge*);

i = *find-index nid* (*inputs-of* (*kind g merge*));
phis = (*phi-list g merge*);
inps = (*phi-inputs g i phis*);
g ⊢ *inps* $\simeq_L$ *inpsE*;
[*m*, *p*] ⊢ *inpsE* $\mapsto_L$ *vs*;

m' = *set-phis phis vs m*⟧
⟹ *g*, *p* ⊢ (*nid*, *m*, *h*) → (*merge*, *m'*, *h*) |

NewInstanceNode:
⟦*kind g nid* = (*NewInstanceNode nid f obj nid'*);
(*h'*, *ref*) = *h-new-inst h*;
m' = *m*(*nid* := *ref*)⟧
⟹ *g*, *p* ⊢ (*nid*, *m*, *h*) → (*nid'*, *m'*, *h'*) |

LoadFieldNode:
$[\![kind\ g\ nid = (LoadFieldNode\ nid\ f\ (Some\ obj)\ nid');$
$g \vdash obj \simeq objE;$
$[m,\ p] \vdash objE \mapsto ObjRef\ ref;$
$h\text{-}load\text{-}field\ f\ ref\ h = v;$
$m' = m(nid := v)]\!]$
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m',\ h)\ |$

SignedDivNode:
$[\![kind\ g\ nid = (SignedDivNode\ nid\ x\ y\ zero\ sb\ nxt);$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye;$
$[m,\ p] \vdash xe \mapsto v1;$
$[m,\ p] \vdash ye \mapsto v2;$
$v = (intval\text{-}div\ v1\ v2);$
$m' = m(nid := v)]\!]$
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \rightarrow (nxt,\ m',\ h)\ |$

SignedRemNode:
$[\![kind\ g\ nid = (SignedRemNode\ nid\ x\ y\ zero\ sb\ nxt);$
$g \vdash x \simeq xe;$
$g \vdash y \simeq ye;$
$[m,\ p] \vdash xe \mapsto v1;$
$[m,\ p] \vdash ye \mapsto v2;$
$v = (intval\text{-}mod\ v1\ v2);$
$m' = m(nid := v)]\!]$
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \rightarrow (nxt,\ m',\ h)\ |$

StaticLoadFieldNode:
$[\![kind\ g\ nid = (LoadFieldNode\ nid\ f\ None\ nid');$
$h\text{-}load\text{-}field\ f\ None\ h = v;$
$m' = m(nid := v)]\!]$
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m',\ h)\ |$

StoreFieldNode:
$[\![kind\ g\ nid = (StoreFieldNode\ nid\ f\ newval\ \text{-}\ (Some\ obj)\ nid');$
$g \vdash newval \simeq newvalE;$
$g \vdash obj \simeq objE;$
$[m,\ p] \vdash newvalE \mapsto val;$
$[m,\ p] \vdash objE \mapsto ObjRef\ ref;$
$h' = h\text{-}store\text{-}field\ f\ ref\ val\ h;$
$m' = m(nid := val)]\!]$
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m',\ h')\ |$

StaticStoreFieldNode:
$[\![kind\ g\ nid = (StoreFieldNode\ nid\ f\ newval\ \text{-}\ None\ nid');$
$g \vdash newval \simeq newvalE;$
$[m,\ p] \vdash newvalE \mapsto val;$

$h' = h\text{-}store\text{-}field\ f\ None\ val\ h$;
$m' = \ m(nid := val)$⟧
$\Longrightarrow g,\ p \vdash (nid,\ m,\ h) \to (nid',\ m',\ h')$

code-pred (*modes*: $i \Rightarrow i \Rightarrow i * i * i \Rightarrow o * o * o \Rightarrow bool$) *step* ⟨*proof*⟩

7.3 Interprocedural Semantics

type-synonym *Signature* = *string*
type-synonym *Program* = *Signature* ⇀ *IRGraph*

inductive *step-top* :: *Program* ⇒ (*IRGraph* × *ID* × *MapState* × *Params*) *list* × *FieldRefHeap* ⇒ (*IRGraph* × *ID* × *MapState* × *Params*) *list* × *FieldRefHeap* ⇒ *bool*
(- ⊢ - ⟶ - *55*)
for *P* **where**

Lift:
⟦$g,\ p \vdash (nid,\ m,\ h) \to (nid',\ m',\ h')$⟧
$\Longrightarrow P \vdash ((g,nid,m,p)\#stk,\ h) \longrightarrow ((g,nid',m',p)\#stk,\ h')$ |

InvokeNodeStep:
⟦*is-Invoke* (*kind g nid*);

callTarget = *ir-callTarget* (*kind g nid*);
kind g callTarget = (*MethodCallTargetNode targetMethod arguments*);
Some targetGraph = *P targetMethod*;
m' = *new-map-state*;
$g \vdash arguments \simeq_L argsE$;
$[m,\ p] \vdash argsE\ \mapsto_L p'$⟧
$\Longrightarrow P \vdash ((g,nid,m,p)\#stk,\ h) \longrightarrow ((targetGraph,0,m',p')\#(g,nid,m,p)\#stk,\ h)$
|

ReturnNode:
⟦*kind g nid* = (*ReturnNode* (*Some expr*) -);
$g \vdash expr \simeq e$;
$[m,\ p] \vdash e \mapsto v$;

$cm' = cm(cnid := v)$;
$cnid'$ = (*successors-of* (*kind cg cnid*))!*0*⟧
$\Longrightarrow P \vdash ((g,nid,m,p)\#(cg,cnid,cm,cp)\#stk,\ h) \longrightarrow ((cg,cnid',cm',cp)\#stk,\ h)$ |

ReturnNodeVoid:
⟦*kind g nid* = (*ReturnNode None* -);
$cm' = cm(cnid := (ObjRef\ (Some\ (2048))))$;

$cnid'$ = (*successors-of* (*kind cg cnid*))!*0*⟧
$\Longrightarrow P \vdash ((g,nid,m,p)\#(cg,cnid,cm,cp)\#stk,\ h) \longrightarrow ((cg,cnid',cm',cp)\#stk,\ h)$ |

UnwindNode:
$[\![$*kind g nid* = (*UnwindNode exception*);

g ⊢ *exception* ≃ *exceptionE*;
[*m*, *p*] ⊢ *exceptionE* ↦ *e*;

kind cg cnid = (*InvokeWithExceptionNode* - - - - - - *exEdge*);

cm′ = *cm*(*cnid* := *e*)$]\!]$
⟹ *P* ⊢ ((*g*,*nid*,*m*,*p*)#(*cg*,*cnid*,*cm*,*cp*)#*stk*, *h*) ⟶ ((*cg*,*exEdge*,*cm′*,*cp*)#*stk*, *h*)

code-pred (*modes*: *i* ⇒ *i* ⇒ *o* ⇒ *bool*) *step-top* ⟨*proof*⟩

7.4 Big-step Execution

type-synonym *Trace* = (*IRGraph* × *ID* × *MapState* × *Params*) *list*

fun *has-return* :: *MapState* ⇒ *bool* **where**
has-return m = (*m 0* ≠ *UndefVal*)

inductive *exec* :: *Program*
⇒ (*IRGraph* × *ID* × *MapState* × *Params*) *list* × *FieldRefHeap*
⇒ *Trace*
⇒ (*IRGraph* × *ID* × *MapState* × *Params*) *list* × *FieldRefHeap*
⇒ *Trace*
⇒ *bool*
(- ⊢ - | - ⟶* - | -)
for *P*
where
$[\![$*P* ⊢ (((*g*,*nid*,*m*,*p*)#*xs*),*h*) ⟶ (((*g′*,*nid′*,*m′*,*p′*)#*ys*),*h′*);
¬(*has-return m′*);

l′ = (*l* @ [(*g*,*nid*,*m*,*p*)]);

exec P (((*g′*,*nid′*,*m′*,*p′*)#*ys*),*h′*) *l′ next-state l″*$]\!]$
⟹ *exec P* (((*g*,*nid*,*m*,*p*)#*xs*),*h*) *l next-state l″*

|
$[\![$*P* ⊢ (((*g*,*nid*,*m*,*p*)#*xs*),*h*) ⟶ (((*g′*,*nid′*,*m′*,*p′*)#*ys*),*h′*);
has-return m′;

l′ = (*l* @ [(*g*,*nid*,*m*,*p*)])$]\!]$
⟹ *exec P* (((*g*,*nid*,*m*,*p*)#*xs*),*h*) *l* (((*g′*,*nid′*,*m′*,*p′*)#*ys*),*h′*) *l′*
code-pred (*modes*: *i* ⇒ *i* ⇒ *i* ⇒ *o* ⇒ *o* ⇒ *bool as Exec*) *exec* ⟨*proof*⟩

inductive *exec-debug* :: *Program*
⇒ (*IRGraph* × *ID* × *MapState* × *Params*) *list* × *FieldRefHeap*

$\Rightarrow$ *nat*
$\Rightarrow$ (*IRGraph* $\times$ *ID* $\times$ *MapState* $\times$ *Params*) *list* $\times$ *FieldRefHeap*
$\Rightarrow$ *bool*
(-$\vdash$-$\longrightarrow$$*$-$*$ -)
where
$[\![n > 0;$
$p \vdash s \longrightarrow s'$;
exec-debug $p\ s'\ (n - 1)\ s'']\!]$
$\Longrightarrow$ *exec-debug* $p\ s\ n\ s''$ |

$[\![n = 0]\!]$
$\Longrightarrow$ *exec-debug* $p\ s\ n\ s$
code-pred (*modes*: $i \Rightarrow i \Rightarrow i \Rightarrow o \Rightarrow$ *bool*) *exec-debug* $\langle$*proof*$\rangle$

7.4.1 Heap Testing

definition *p3*:: *Params* **where**
p3 = [*IntVal 32 3*]

values {(*prod.fst*(*prod.snd* (*prod.snd* (*hd* (*prod.fst res*))))) *0*
| *res*. (λx . *Some eg2-sq*) $\vdash$ ([(*eg2-sq*,*0*,*new-map-state*,*p3*), (*eg2-sq*,*0*,*new-map-state*,*p3*)], *new-heap*) $\rightarrow$$*$*2*$*$ *res*}

definition *field-sq* :: *string* **where**
field-sq = ''*sq*''

definition *eg3-sq* :: *IRGraph* **where**
eg3-sq = *irgraph* [
(*0*, *StartNode None 4*, *VoidStamp*),
(*1*, *ParameterNode 0*, *default-stamp*),
(*3*, *MulNode 1 1*, *default-stamp*),
(*4*, *StoreFieldNode 4 field-sq 3 None None 5*, *VoidStamp*),
(*5*, *ReturnNode* (*Some 3*) *None*, *default-stamp*)
]

values {*h-load-field field-sq None* (*prod.snd res*)
| *res*. (λx. *Some eg3-sq*) $\vdash$ ([(*eg3-sq*, *0*, *new-map-state*, *p3*), (*eg3-sq*, *0*, *new-map-state*, *p3*)], *new-heap*) $\rightarrow$$*$*3*$*$ *res*}

definition *eg4-sq* :: *IRGraph* **where**
eg4-sq = *irgraph* [
(*0*, *StartNode None 4*, *VoidStamp*),
(*1*, *ParameterNode 0*, *default-stamp*),
(*3*, *MulNode 1 1*, *default-stamp*),
(*4*, *NewInstanceNode 4* ''*obj-class*'' *None 5*, *ObjectStamp* ''*obj-class*'' *True True True*),
(*5*, *StoreFieldNode 5 field-sq 3 None* (*Some 4*) *6*, *VoidStamp*),

(*6*, *ReturnNode* (*Some 3*) *None*, *default-stamp*)
]

values {*h-load-field field-sq* (*Some 0*) (*prod.snd res*) | *res*.
(*λx. Some eg4-sq*) ⊢ ([(*eg4-sq*, *0*, *new-map-state*, *p3*), (*eg4-sq*, *0*, *new-map-state*, *p3*)], *new-heap*) →∗*3*∗ *res*}

end

7.5 Control-flow Semantics Theorems

theory *IRStepThms*
imports
IRStepObj
TreeToGraphThms
begin

We prove that within the same graph, a configuration triple will always transition to the same subsequent configuration. Therefore, our step semantics is deterministic.

7.5.1 Control-flow Step is Deterministic

theorem *stepDet*:
$(g,\ p \vdash (nid,m,h) \rightarrow next) \Longrightarrow$
$(\forall\ next'.\ ((g,\ p \vdash (nid,m,h) \rightarrow next') \longrightarrow next = next'))$
⟨*proof*⟩

lemma *stepRefNode*:
$[\![kind\ g\ nid = RefNode\ nid']\!] \Longrightarrow g,\ p \vdash (nid,m,h) \rightarrow (nid',m,h)$
⟨*proof*⟩

lemma *IfNodeStepCases*:
assumes $kind\ g\ nid = IfNode\ cond\ tb\ fb$
assumes $g \vdash cond \simeq condE$
assumes $[m,\ p] \vdash condE \mapsto v$
assumes $g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m,\ h)$
shows $nid' \in \{tb,\ fb\}$
⟨*proof*⟩

lemma *IfNodeSeq*:
shows $kind\ g\ nid = IfNode\ cond\ tb\ fb \longrightarrow \neg(is\text{-}sequential\text{-}node\ (kind\ g\ nid))$
⟨*proof*⟩

lemma *IfNodeCond*:
assumes $kind\ g\ nid = IfNode\ cond\ tb\ fb$
assumes $g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m,\ h)$
shows $\exists\ condE\ v.\ ((g \vdash cond \simeq condE) \wedge ([m,\ p] \vdash condE \mapsto v))$

$\langle proof \rangle$

lemma *step-in-ids*:
assumes $g,\ p \vdash (nid,\ m,\ h) \to (nid',\ m',\ h')$
shows $nid \in ids\ g$
$\langle proof \rangle$

end

8 Proof Infrastructure

8.1 Bisimulation

theory *Bisimulation*
imports
Stuttering
begin

inductive *weak-bisimilar* :: $ID \Rightarrow IRGraph \Rightarrow IRGraph \Rightarrow bool$
(- . - $\sim$ -) **for** *nid* **where**
$[\![\forall P'.\ (g\ m\ p\ h \vdash nid \leadsto P') \longrightarrow (\exists\, Q'\ .\ (g'\ m\ p\ h \vdash nid \leadsto Q') \land P' = Q');$
$\forall Q'.\ (g'\ m\ p\ h \vdash nid \leadsto Q') \longrightarrow (\exists\, P'\ .\ (g\ m\ p\ h \vdash nid \leadsto P') \land P' = Q')]\!]$
$\Longrightarrow nid\ .\ g \sim g'$

A strong bisimilution between no-op transitions

inductive *strong-noop-bisimilar* :: $ID \Rightarrow IRGraph \Rightarrow IRGraph \Rightarrow bool$
(- | - $\sim$ -) **for** *nid* **where**
$[\![\forall P'.\ (g,\ p \vdash (nid,\ m,\ h) \to P') \longrightarrow (\exists\, Q'\ .\ (g',\ p \vdash (nid,\ m,\ h) \to Q') \land P' = Q');$
$\forall Q'.\ (g',\ p \vdash (nid,\ m,\ h) \to Q') \longrightarrow (\exists\, P'\ .\ (g,\ p \vdash (nid,\ m,\ h) \to P') \land P' = Q')]\!]$
$\Longrightarrow nid \mid g \sim g'$

lemma *lockstep-strong-bisimilulation*:
assumes $g' = replace\text{-}node\ nid\ node\ g$
assumes $g,\ p \vdash (nid,\ m,\ h) \to (nid',\ m,\ h)$
assumes $g',\ p \vdash (nid,\ m,\ h) \to (nid',\ m,\ h)$
shows $nid \mid g \sim g'$
$\langle proof \rangle$

lemma *no-step-bisimulation*:
assumes $\forall\, m\ p\ h\ nid'\ m'\ h'.\ \neg(g,\ p \vdash (nid,\ m,\ h) \to (nid',\ m',\ h'))$
assumes $\forall\, m\ p\ h\ nid'\ m'\ h'.\ \neg(g',\ p \vdash (nid,\ m,\ h) \to (nid',\ m',\ h'))$
shows $nid \mid g \sim g'$
$\langle proof \rangle$

end

8.2 Graph Rewriting

theory
  Rewrites
imports
  Stuttering
begin

fun *replace-usages* :: $ID \Rightarrow ID \Rightarrow IRGraph \Rightarrow IRGraph$ **where**
  replace-usages nid nid′ g = replace-node nid (RefNode nid′, stamp g nid′) g

lemma *replace-usages-effect*:
  assumes *g′ = replace-usages nid nid′ g*
  shows *kind g′ nid = RefNode nid′*
  ⟨*proof*⟩

lemma *replace-usages-changeonly*:
  assumes $nid \in ids\ g$
  assumes *g′ = replace-usages nid nid′ g*
  shows *changeonly* $\{nid\}$ *g g′*
  ⟨*proof*⟩

lemma *replace-usages-unchanged*:
  assumes $nid \in ids\ g$
  assumes *g′ = replace-usages nid nid′ g*
  shows *unchanged* $(ids\ g - \{nid\})$ *g g′*
  ⟨*proof*⟩

fun *nextNid* :: $IRGraph \Rightarrow ID$ **where**
  nextNid g = (Max (ids g)) + 1

lemma *max-plus-one*:
  fixes *c* :: *ID set*
  shows $[\![finite\ c;\ c \neq \{\}]\!] \Longrightarrow (Max\ c) + 1 \notin c$
  ⟨*proof*⟩

lemma *ids-finite*:
  finite (ids g)
  ⟨*proof*⟩

lemma *nextNidNotIn*:
  $ids\ g \neq \{\} \longrightarrow nextNid\ g \notin ids\ g$
  ⟨*proof*⟩

fun *bool-to-val-width1* :: $bool \Rightarrow Value$ **where**
  bool-to-val-width1 True = (IntVal 1 1) |
  bool-to-val-width1 False = (IntVal 1 0)

fun *constantCondition* :: *bool* ⇒ *ID* ⇒ *IRNode* ⇒ *IRGraph* ⇒ *IRGraph* **where**
constantCondition val nid (*IfNode cond t f*) *g* =
replace-node nid (*IfNode* (*nextNid g*) *t f*, *stamp g nid*)
(*add-node* (*nextNid g*) ((*ConstantNode* (*bool-to-val-width1 val*)), *constantAsStamp* (*bool-to-val-width1 val*)) *g*) |
constantCondition cond nid - g = *g*

lemma *constantConditionTrue*:
assumes *kind g ifcond* = *IfNode cond t f*
assumes g' = *constantCondition True ifcond* (*kind g ifcond*) *g*
shows g', $p \vdash (ifcond,\ m,\ h) \rightarrow (t,\ m,\ h)$
⟨*proof*⟩

lemma *constantConditionFalse*:
assumes *kind g ifcond* = *IfNode cond t f*
assumes g' = *constantCondition False ifcond* (*kind g ifcond*) *g*
shows g', $p \vdash (ifcond,\ m,\ h) \rightarrow (f,\ m,\ h)$
⟨*proof*⟩

lemma *diff-forall*:
assumes $\forall\, n \in ids\ g - \{nid\}.\ cond\ n$
shows $\forall\, n.\ n \in ids\ g \land n \notin \{nid\} \longrightarrow cond\ n$
⟨*proof*⟩

lemma *replace-node-changeonly*:
assumes g' = *replace-node nid node g*
shows *changeonly* {*nid*} *g* g'
⟨*proof*⟩

lemma *add-node-changeonly*:
assumes g' = *add-node nid node g*
shows *changeonly* {*nid*} *g* g'
⟨*proof*⟩

lemma *constantConditionNoEffect*:
assumes ¬(*is-IfNode* (*kind g nid*))
shows *g* = *constantCondition b nid* (*kind g nid*) *g*
⟨*proof*⟩

lemma *constantConditionIfNode*:
assumes *kind g nid* = *IfNode cond t f*
shows *constantCondition val nid* (*kind g nid*) *g* =
replace-node nid (*IfNode* (*nextNid g*) *t f*, *stamp g nid*)
(*add-node* (*nextNid g*) ((*ConstantNode* (*bool-to-val-width1 val*)), *constantAsStamp* (*bool-to-val-width1 val*)) *g*)
⟨*proof*⟩

lemma *constantCondition-changeonly*:
assumes $nid \in ids\ g$

assumes $g' = constantCondition\ b\ nid\ (kind\ g\ nid)\ g$
shows $changeonly\ \{nid\}\ g\ g'$
⟨*proof*⟩

lemma *constantConditionNoIf*:
assumes $\forall\, cond\ t\ f.\ kind\ g\ ifcond \neq IfNode\ cond\ t\ f$
assumes $g' = constantCondition\ val\ ifcond\ (kind\ g\ ifcond)\ g$
shows $\exists\, nid'\ .(g\ m\ p\ h \vdash ifcond \leadsto nid') \longleftrightarrow (g'\ m\ p\ h \vdash ifcond \leadsto nid')$
⟨*proof*⟩

lemma *constantConditionValid*:
assumes $kind\ g\ ifcond = IfNode\ cond\ t\ f$
assumes $[g,\ m,\ p] \vdash cond \mapsto v$
assumes $const = val\text{-}to\text{-}bool\ v$
assumes $g' = constantCondition\ const\ ifcond\ (kind\ g\ ifcond)\ g$
shows $\exists\, nid'\ .(g\ m\ p\ h \vdash ifcond \leadsto nid') \longleftrightarrow (g'\ m\ p\ h \vdash ifcond \leadsto nid')$
⟨*proof*⟩

end

8.3 Stuttering

theory *Stuttering*
imports
Semantics.IRStepThms
begin

inductive *stutter*:: $IRGraph \Rightarrow MapState \Rightarrow Params \Rightarrow FieldRefHeap \Rightarrow ID \Rightarrow ID \Rightarrow bool$ (- - - - ⊢ - ⇝ - 55)
for $g\ m\ p\ h$ **where**

StutterStep:
$[\![g,\ p \vdash (nid,m,h) \rightarrow (nid',m,h)]\!]$
$\Longrightarrow g\ m\ p\ h \vdash nid \leadsto nid'$ |

Transitive:
$[\![g,\ p \vdash (nid,m,h) \rightarrow (nid'',m,h);$
$g\ m\ p\ h \vdash nid'' \leadsto nid']\!]$
$\Longrightarrow g\ m\ p\ h \vdash nid \leadsto nid'$

lemma *stuttering-successor*:
assumes $(g,\ p \vdash (nid,\ m,\ h) \rightarrow (nid',\ m,\ h))$
shows $\{P'.\ (g\ m\ p\ h \vdash nid \leadsto P')\} = \{nid'\} \cup \{nid''.\ (g\ m\ p\ h \vdash nid' \leadsto nid'')\}$
⟨*proof*⟩

end

8.4 Evaluation Stamp Theorems

theory *StampEvalThms*
  imports *Graph.ValueThms*
    Semantics.IRTreeEvalThms
begin

lemma
  assumes *take-bit b v = v*
  shows *signed-take-bit b v = v*
  ⟨*proof*⟩

lemma *unwrap-signed-take-bit*:
  fixes *v :: int64*
  assumes *0 < b ∧ b ≤ 64*
  assumes *signed-take-bit (b − 1) v = v*
  shows *signed-take-bit 63 (Word.rep (signed-take-bit (b − Suc 0) v)) = sint v*
  ⟨*proof*⟩

lemma *unrestricted-new-int-always-valid* [*simp*]:
  assumes *0 < b ∧ b ≤ 64*
  shows *valid-value (new-int b v) (unrestricted-stamp (IntegerStamp b lo hi))*
  ⟨*proof*⟩

lemma *unary-undef*: *val = UndefVal ⟹ unary-eval op val = UndefVal*
  ⟨*proof*⟩

lemma *unary-obj*: *val = ObjRef x ⟹ unary-eval op val = UndefVal*
  ⟨*proof*⟩

lemma *unrestricted-stamp-valid*:
  assumes *s = unrestricted-stamp (IntegerStamp b lo hi)*
  assumes *0 < b ∧ b ≤ 64*
  shows *valid-stamp s*
  ⟨*proof*⟩

lemma *unrestricted-stamp-valid-value* [*simp*]:
  assumes *1*: *result = IntVal b ival*
  assumes *take-bit b ival = ival*
  assumes *0 < b ∧ b ≤ 64*
  shows *valid-value result (unrestricted-stamp (IntegerStamp b lo hi))*
⟨*proof*⟩

8.4.1 Support Lemmas for Integer Stamps and Associated IntVal values

Valid int implies some useful facts.

lemma *valid-int-gives*:
assumes *valid-value* (*IntVal b val*) *stamp*
obtains *lo hi* **where** *stamp* = *IntegerStamp b lo hi* ∧
valid-stamp (*IntegerStamp b lo hi*) ∧
take-bit b val = *val* ∧
lo ≤ *int-signed-value b val* ∧ *int-signed-value b val* ≤ *hi*
⟨*proof*⟩

And the corresponding lemma where we know the stamp rather than the value.

lemma *valid-int-stamp-gives*:
assumes *valid-value val* (*IntegerStamp b lo hi*)
obtains *ival* **where** *val* = *IntVal b ival* ∧
valid-stamp (*IntegerStamp b lo hi*) ∧
take-bit b ival = *ival* ∧
lo ≤ *int-signed-value b ival* ∧ *int-signed-value b ival* ≤ *hi*
⟨*proof*⟩

A valid int must have the expected number of bits.

lemma *valid-int-same-bits*:
assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
shows *b* = *bits*
⟨*proof*⟩

A valid value means a valid stamp.

lemma *valid-int-valid-stamp*:
assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
shows *valid-stamp* (*IntegerStamp bits lo hi*)
⟨*proof*⟩

A valid int means a valid non-empty stamp.

lemma *valid-int-not-empty*:
assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
shows *lo* ≤ *hi*
⟨*proof*⟩

A valid int fits into the given number of bits (and other bits are zero).

lemma *valid-int-fits*:
assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
shows *take-bit bits val* = *val*
⟨*proof*⟩

lemma *valid-int-is-zero-masked*:
assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
shows *and val* (*not* (*mask bits*)) = *0*
⟨*proof*⟩

Unsigned ints have bounds 0 up to 2^bits.

lemma *valid-int-unsigned-bounds*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)

  shows *uint val* $<$ *2* ^ *bits*
  ⟨*proof*⟩

Signed ints have the usual two-complement bounds.

lemma *valid-int-signed-upper-bound*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
  shows *int-signed-value bits val* $<$ *2* ^(*bits* $-$ *1*)
  ⟨*proof*⟩

lemma *valid-int-signed-lower-bound*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
  shows $-$(*2* ^(*bits* $-$ *1*)) $\leq$ *int-signed-value bits val*
  ⟨*proof*⟩

and *bit_bounds* versions of the above bounds.

lemma *valid-int-signed-upper-bit-bound*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
  shows *int-signed-value bits val* $\leq$ *snd* (*bit-bounds bits*)
⟨*proof*⟩

lemma *valid-int-signed-lower-bit-bound*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
  shows *fst* (*bit-bounds bits*) $\leq$ *int-signed-value bits val*
⟨*proof*⟩

Valid values satisfy their stamp bounds.

lemma *valid-int-signed-range*:
  assumes *valid-value* (*IntVal b val*) (*IntegerStamp bits lo hi*)
  shows *lo* $\leq$ *int-signed-value bits val* $\wedge$ *int-signed-value bits val* $\leq$ *hi*
  ⟨*proof*⟩

8.4.2 Validity of all Unary Operators

We split the validity proof for unary operators into two lemmas, one for normal unary operators whose output bits equals their input bits, and the other case for the widen and narrow operators.

lemma *eval-normal-unary-implies-valid-value*:
  assumes [*m*,*p*] $\vdash$ *expr* $\mapsto$ *val*
  assumes *result* $=$ *unary-eval op val*
  assumes *op*: *op* $\in$ *normal-unary*
  assumes *result* $\neq$ *UndefVal*
  assumes *valid-value val* (*stamp-expr expr*)
  shows *valid-value result* (*stamp-expr* (*UnaryExpr op expr*))
⟨*proof*⟩

lemma *narrow-widen-output-bits*:
  assumes *unary-eval op val* $\neq$ *UndefVal*
  assumes *op* $\notin$ *normal-unary*
  shows $0 < (\textit{ir-resultBits op}) \wedge (\textit{ir-resultBits op}) \leq 64$
$\langle proof \rangle$

lemma *eval-widen-narrow-unary-implies-valid-value*:
  assumes $[m,p] \vdash \textit{expr} \mapsto \textit{val}$
  assumes *result* = *unary-eval op val*
  assumes *op*: *op* $\notin$ *normal-unary*
  assumes *result* $\neq$ *UndefVal*
  assumes *valid-value val* (*stamp-expr expr*)
  shows *valid-value result* (*stamp-expr* (*UnaryExpr op expr*))
$\langle proof \rangle$

lemma *eval-unary-implies-valid-value*:
  assumes $[m,p] \vdash \textit{expr} \mapsto \textit{val}$
  assumes *result* = *unary-eval op val*
  assumes *result* $\neq$ *UndefVal*
  assumes *valid-value val* (*stamp-expr expr*)
  shows *valid-value result* (*stamp-expr* (*UnaryExpr op expr*))
  $\langle proof \rangle$

8.4.3 Support Lemmas for Binary Operators

lemma *binary-undef*: *v1* = *UndefVal* $\vee$ *v2* = *UndefVal* $\Longrightarrow$ *bin-eval op v1 v2* = *UndefVal*
  $\langle proof \rangle$

lemma *binary-obj*: *v1* = *ObjRef x* $\vee$ *v2* = *ObjRef y* $\Longrightarrow$ *bin-eval op v1 v2* = *UndefVal*
  $\langle proof \rangle$

Some lemmas about the three different output sizes for binary operators.

lemma *bin-eval-bits-binary-shift-ops*:
  assumes *result* = *bin-eval op* (*IntVal b1 v1*) (*IntVal b2 v2*)
  assumes *result* $\neq$ *UndefVal*
  assumes *op* $\in$ *binary-shift-ops*
  shows $\exists$ *v. result* = *new-int b1 v*
  $\langle proof \rangle$

lemma *bin-eval-bits-fixed-32-ops*:
  assumes *result* = *bin-eval op* (*IntVal b1 v1*) (*IntVal b2 v2*)
  assumes *result* $\neq$ *UndefVal*
  assumes *op* $\in$ *binary-fixed-32-ops*
  shows $\exists$ *v. result* = *new-int 32 v*
  $\langle proof \rangle$

lemma *bin-eval-bits-normal-ops*:
assumes *result* = *bin-eval op* (*IntVal b1 v1*) (*IntVal b2 v2*)
assumes *result* ≠ *UndefVal*
assumes *op* ∉ *binary-shift-ops*
assumes *op* ∉ *binary-fixed-32-ops*
shows ∃ *v*. *result* = *new-int b1 v*
⟨*proof*⟩

lemma *bin-eval-input-bits-equal*:
assumes *result* = *bin-eval op* (*IntVal b1 v1*) (*IntVal b2 v2*)
assumes *result* ≠ *UndefVal*
assumes *op* ∉ *binary-shift-ops*
shows *b1* = *b2*
⟨*proof*⟩

lemma *bin-eval-implies-valid-value*:
assumes [*m*,*p*] ⊢ *expr1* ↦ *val1*
assumes [*m*,*p*] ⊢ *expr2* ↦ *val2*
assumes *result* = *bin-eval op val1 val2*
assumes *result* ≠ *UndefVal*
assumes *valid-value val1* (*stamp-expr expr1*)
assumes *valid-value val2* (*stamp-expr expr2*)
shows *valid-value result* (*stamp-expr* (*BinaryExpr op expr1 expr2*))
⟨*proof*⟩

8.4.4 Validity of Stamp Meet and Join Operators

lemma *stamp-meet-integer-is-valid-stamp*:
assumes *valid-stamp stamp1*
assumes *valid-stamp stamp2*
assumes *is-IntegerStamp stamp1*
assumes *is-IntegerStamp stamp2*
shows *valid-stamp* (*meet stamp1 stamp2*)
⟨*proof*⟩

lemma *stamp-meet-is-valid-stamp*:
assumes *1*: *valid-stamp stamp1*
assumes *2*: *valid-stamp stamp2*
shows *valid-stamp* (*meet stamp1 stamp2*)
⟨*proof*⟩

lemma *stamp-meet-commutes*: *meet stamp1 stamp2* = *meet stamp2 stamp1*
⟨*proof*⟩

lemma *stamp-meet-is-valid-value1*:
assumes *valid-value val stamp1*
assumes *valid-stamp stamp2*

assumes *stamp1* = *IntegerStamp b1 lo1 hi1*
assumes *stamp2* = *IntegerStamp b2 lo2 hi2*
assumes *meet stamp1 stamp2* ≠ *IllegalStamp*
shows *valid-value val* (*meet stamp1 stamp2*)
⟨*proof*⟩

and the symmetric lemma follows by the commutativity of meet.

lemma *stamp-meet-is-valid-value*:
assumes *valid-value val stamp2*
assumes *valid-stamp stamp1*
assumes *stamp1* = *IntegerStamp b1 lo1 hi1*
assumes *stamp2* = *IntegerStamp b2 lo2 hi2*
assumes *meet stamp1 stamp2* ≠ *IllegalStamp*
shows *valid-value val* (*meet stamp1 stamp2*)
⟨*proof*⟩

8.4.5 Validity of conditional expressions

lemma *conditional-eval-implies-valid-value*:
assumes [*m*,*p*] ⊢ *cond* ↦ *condv*
assumes *expr* = (*if val-to-bool condv then expr1 else expr2*)
assumes [*m*,*p*] ⊢ *expr* ↦ *val*
assumes *val* ≠ *UndefVal*
assumes *valid-value condv* (*stamp-expr cond*)
assumes *valid-value val* (*stamp-expr expr*)
assumes *compatible* (*stamp-expr expr1*) (*stamp-expr expr2*)
shows *valid-value val* (*stamp-expr* (*ConditionalExpr cond expr1 expr2*))
⟨*proof*⟩

8.4.6 Validity of Whole Expression Tree Evaluation

TODO: find a way to encode that conditional expressions must have compatible (and valid) stamps? One approach would be for all the stamp_expr operators to require that all input stamps are valid.

definition *wf-stamp* :: *IRExpr* ⇒ *bool* **where**
wf-stamp e = (∀ *m p v*. ([*m*, *p*] ⊢ *e* ↦ *v*) ⟶ *valid-value v* (*stamp-expr e*))

lemma *stamp-under-defn*:
assumes *stamp-under* (*stamp-expr x*) (*stamp-expr y*)
assumes *wf-stamp x* ∧ *wf-stamp y*
assumes ([*m*, *p*] ⊢ *x* ↦ *xv*) ∧ ([*m*, *p*] ⊢ *y* ↦ *yv*)
shows *val-to-bool* (*bin-eval BinIntegerLessThan xv yv*) ∨ (*bin-eval BinIntegerLessThan xv yv*) = *UndefVal*
⟨*proof*⟩

lemma *stamp-under-defn-inverse*:
assumes *stamp-under* (*stamp-expr y*) (*stamp-expr x*)
assumes *wf-stamp x* ∧ *wf-stamp y*

assumes $([m, p] \vdash x \mapsto xv) \wedge ([m, p] \vdash y \mapsto yv)$
shows $\neg$(*val-to-bool* (*bin-eval BinIntegerLessThan xv yv*)) $\vee$ (*bin-eval BinIntegerLessThan xv yv*) = *UndefVal*
⟨*proof*⟩

end

9 Optization DSL

9.1 Markup

theory *Markup*
imports *Semantics.IRTreeEval Snippets.Snipping*
begin

datatype *'a Rewrite* =
Transform 'a 'a (- $\longmapsto$ - *10*) |
Conditional 'a 'a bool (- $\longmapsto$ - *when* - *11*) |
Sequential 'a Rewrite 'a Rewrite |
Transitive 'a Rewrite

datatype *'a ExtraNotation* =
ConditionalNotation 'a 'a 'a (- *?* - : - *50*) |
EqualsNotation 'a 'a (- *eq* -) |
ConstantNotation 'a (*const* - *120*) |
TrueNotation (*true*) |
FalseNotation (*false*) |
ExclusiveOr 'a 'a (- $\oplus$ -) |
LogicNegationNotation 'a (!-) |
ShortCircuitOr 'a 'a (- || -)

definition *word* :: (*'a*::*len*) *word* $\Rightarrow$ *'a word* **where**
word x = *x*

ML-file ‹*markup.ML*›

9.1.1 Expression Markup

ML ‹
structure IRExprTranslator : *DSL-TRANSLATION* =
struct
fun markup DSL-Tokens.Add = @{*term BinaryExpr*} \$ @{*term BinAdd*}
| *markup DSL-Tokens.Sub* = @{*term BinaryExpr*} \$ @{*term BinSub*}
| *markup DSL-Tokens.Mul* = @{*term BinaryExpr*} \$ @{*term BinMul*}
| *markup DSL-Tokens.And* = @{*term BinaryExpr*} \$ @{*term BinAnd*}
| *markup DSL-Tokens.Or* = @{*term BinaryExpr*} \$ @{*term BinOr*}
| *markup DSL-Tokens.Xor* = @{*term BinaryExpr*} \$ @{*term BinXor*}
| *markup DSL-Tokens.ShortCircuitOr* = @{*term BinaryExpr*} \$ @{*term BinShortCircuitOr*}

| *markup DSL-Tokens.Abs* = @{*term UnaryExpr*} $ @{*term UnaryAbs*}
| *markup DSL-Tokens.Less* = @{*term BinaryExpr*} $ @{*term BinIntegerLessThan*}
| *markup DSL-Tokens.Equals* = @{*term BinaryExpr*} $ @{*term BinIntegerEquals*}
| *markup DSL-Tokens.Not* = @{*term UnaryExpr*} $ @{*term UnaryNot*}
| *markup DSL-Tokens.Negate* = @{*term UnaryExpr*} $ @{*term UnaryNeg*}
| *markup DSL-Tokens.LogicNegate* = @{*term UnaryExpr*} $ @{*term UnaryLogicNegation*}
| *markup DSL-Tokens.LeftShift* = @{*term BinaryExpr*} $ @{*term BinLeftShift*}
| *markup DSL-Tokens.RightShift* = @{*term BinaryExpr*} $ @{*term BinRightShift*}
| *markup DSL-Tokens.UnsignedRightShift* = @{*term BinaryExpr*} $ @{*term BinURightShift*}
| *markup DSL-Tokens.Conditional* = @{*term ConditionalExpr*}
| *markup DSL-Tokens.Constant* = @{*term ConstantExpr*}
| *markup DSL-Tokens.TrueConstant* = @{*term ConstantExpr* (*IntVal 32 1*)}
| *markup DSL-Tokens.FalseConstant* = @{*term ConstantExpr* (*IntVal 32 0*)}
end
structure IRExprMarkup = *DSL-Markup*(*IRExprTranslator*);
›

ir expression translation

syntax *-expandExpr* :: *term* ⇒ *term* (*exp*[-])
parse-translation ‹ [(@{*syntax-const -expandExpr*} , *IRExprMarkup.markup-expr* [])] ›

ir expression example

value *exp*[(e_1 < e_2) *?* e_1 : e_2]

ConditionalExpr (*BinaryExpr BinIntegerLessThan* e_1 e_2) e_1 e_2

9.1.2 Value Markup

ML ‹
structure IntValTranslator : *DSL-TRANSLATION* =
struct
fun markup DSL-Tokens.Add = @{*term intval-add*}
| *markup DSL-Tokens.Sub* = @{*term intval-sub*}
| *markup DSL-Tokens.Mul* = @{*term intval-mul*}
| *markup DSL-Tokens.And* = @{*term intval-and*}
| *markup DSL-Tokens.Or* = @{*term intval-or*}
| *markup DSL-Tokens.ShortCircuitOr* = @{*term intval-short-circuit-or*}
| *markup DSL-Tokens.Xor* = @{*term intval-xor*}
| *markup DSL-Tokens.Abs* = @{*term intval-abs*}
| *markup DSL-Tokens.Less* = @{*term intval-less-than*}
| *markup DSL-Tokens.Equals* = @{*term intval-equals*}
| *markup DSL-Tokens.Not* = @{*term intval-not*}
| *markup DSL-Tokens.Negate* = @{*term intval-negate*}
| *markup DSL-Tokens.LogicNegate* = @{*term intval-logic-negation*}

| *markup DSL-Tokens.LeftShift* = @{*term intval-left-shift*}
| *markup DSL-Tokens.RightShift* = @{*term intval-right-shift*}
| *markup DSL-Tokens.UnsignedRightShift* = @{*term intval-uright-shift*}
| *markup DSL-Tokens.Conditional* = @{*term intval-conditional*}
| *markup DSL-Tokens.Constant* = @{*term IntVal 32*}
| *markup DSL-Tokens.TrueConstant* = @{*term IntVal 32 1*}
| *markup DSL-Tokens.FalseConstant* = @{*term IntVal 32 0*}
end
structure IntValMarkup = *DSL-Markup*(*IntValTranslator*);
›

value expression translation

syntax *-expandIntVal* :: *term* ⇒ *term* (*val*[-])
parse-translation ‹ [(@{*syntax-const -expandIntVal*} , *IntValMarkup.markup-expr* [])] ›

value expression example

value *val*[(e_1 < e_2) ? e_1 : e_2]

intval-conditional (*intval-less-than* e_1 e_2) e_1 e_2

9.1.3 Word Markup

ML ‹
structure WordTranslator : *DSL-TRANSLATION* =
struct
fun markup DSL-Tokens.Add = @{*term plus*}
| *markup DSL-Tokens.Sub* = @{*term minus*}
| *markup DSL-Tokens.Mul* = @{*term times*}
| *markup DSL-Tokens.And* = @{*term Bit-Operations.semiring-bit-operations-class.and*}
| *markup DSL-Tokens.Or* = @{*term or*}
| *markup DSL-Tokens.Xor* = @{*term xor*}
| *markup DSL-Tokens.Abs* = @{*term abs*}
| *markup DSL-Tokens.Less* = @{*term less*}
| *markup DSL-Tokens.Equals* = @{*term HOL.eq*}
| *markup DSL-Tokens.Not* = @{*term not*}
| *markup DSL-Tokens.Negate* = @{*term uminus*}
| *markup DSL-Tokens.LogicNegate* = @{*term logic-negate*}
| *markup DSL-Tokens.LeftShift* = @{*term shiftl*}
| *markup DSL-Tokens.RightShift* = @{*term signed-shiftr*}
| *markup DSL-Tokens.UnsignedRightShift* = @{*term shiftr*}
| *markup DSL-Tokens.Constant* = @{*term word*}
| *markup DSL-Tokens.TrueConstant* = @{*term 1*}
| *markup DSL-Tokens.FalseConstant* = @{*term 0*}
end
structure WordMarkup = *DSL-Markup*(*WordTranslator*);
›

word expression translation

syntax *-expandWord* :: *term* ⇒ *term* (*bin*[-])
parse-translation ‹ [(@{*syntax-const* *-expandWord*} , *Word-Markup.markup-expr* [])] ›

word expression example

value *bin*[*x* & *y* | *z*]

intval-conditional (*intval-less-than* e_1 e_2) e_1 e_2

value *bin*[$-x$]
value *val*[$-x$]
value *exp*[$-x$]

value *bin*[!*x*]
value *val*[!*x*]
value *exp*[!*x*]

value *bin*[¬*x*]
value *val*[¬*x*]
value *exp*[¬*x*]

value *bin*[~*x*]
value *val*[~*x*]
value *exp*[~*x*]

value ~*x*

end

9.2 Optimization Phases

theory *Phase*
  imports *Main*
begin

ML-file *map.ML*
ML-file *phase.ML*

end

9.3 Canonicalization DSL

theory *Canonicalization*
  imports
    Markup
    Phase

```
    HOL−Eisbach.Eisbach
  keywords
    phase :: thy-decl and
    terminating :: quasi-command and
    print-phases :: diag and
    export-phases :: thy-decl and
    optimization :: thy-goal-defn
begin

print-methods

ML ‹
datatype 'a Rewrite =
  Transform of 'a * 'a |
  Conditional of 'a * 'a * term |
  Sequential of 'a Rewrite * 'a Rewrite |
  Transitive of 'a Rewrite

type rewrite = {
  name: binding,
  rewrite: term Rewrite,
  proofs: thm list,
  code: thm list,
  source: term
}

structure RewriteRule : Rule =
struct
type T = rewrite;

(*
fun pretty-rewrite ctxt (Transform (from, to)) =
      Pretty.block [
        Syntax.pretty-term ctxt from,
        Pretty.str  ↦ ,
        Syntax.pretty-term ctxt to
      ]
  | pretty-rewrite ctxt (Conditional (from, to, cond)) =
      Pretty.block [
        Syntax.pretty-term ctxt from,
        Pretty.str  ↦ ,
        Syntax.pretty-term ctxt to,
        Pretty.str  when ,
        Syntax.pretty-term ctxt cond
      ]
  | pretty-rewrite - - = Pretty.str not implemented*)

fun pretty-thm ctxt thm =
  (Proof-Context.pretty-fact ctxt (, [thm]))
```

```
fun pretty ctxt obligations t =
  let
    val is-skipped = Thm-Deps.has-skip-proof (#proofs t);

    val warning = (if is-skipped
      then [Pretty.str (proof skipped), Pretty.brk 0]
      else []);

    val obligations = (if obligations
      then [Pretty.big-list
              obligations:
              (map (pretty-thm ctxt) (#proofs t)),
            Pretty.brk 0]
      else []);

    fun pretty-bind binding =
      Pretty.markup
        (Position.markup (Binding.pos-of binding) Markup.position)
        [Pretty.str (Binding.name-of binding)];

  in
  Pretty.block ([
    pretty-bind (#name t), Pretty.str : ,
    Syntax.pretty-term ctxt (#source t), Pretty.fbrk
  ] @ obligations @ warning)
  end
end

structure RewritePhase = DSL-Phase(RewriteRule);

val - =
  Outer-Syntax.command command-keyword ‹phase› enter an optimization phase
   (Parse.binding --| Parse.$$$ terminating -- Parse.const --| Parse.begin
     >> (Toplevel.begin-main-target true o RewritePhase.setup));

fun print-phases print-obligations ctxt =
  let
    val thy = Proof-Context.theory-of ctxt;
    fun print phase = RewritePhase.pretty print-obligations phase ctxt
  in
    map print (RewritePhase.phases thy)
  end

fun print-optimizations print-obligations thy =
  print-phases print-obligations thy |> Pretty.writeln-chunks

val - =
  Outer-Syntax.command command-keyword ‹print-phases›
```

```
    print debug information for optimizations
    (Parse.opt-bang >>
      (fn b => Toplevel.keep ((print-optimizations b) o Toplevel.context-of)));

fun export-phases thy name =
  let
    val state = Toplevel.theory-toplevel thy;
    val ctxt = Toplevel.context-of state;
    val content = Pretty.string-of (Pretty.chunks (print-phases false ctxt));
    val cleaned = YXML.content-of content;


    val filename = Path.explode (name^.rules);
    val directory = Path.explode optimizations;
    val path = Path.binding (
                Path.append directory filename,
                Position.none);
    val thy' = thy |> Generated-Files.add-files (path, (Bytes.string content));

    val - = Export.export thy' path [YXML.parse cleaned];

    val - = writeln (Export.message thy' (Path.basic optimizations));
  in
    thy'
  end

val - =
  Outer-Syntax.command command-keyword ‹export-phases›
    export information about encoded optimizations
    (Parse.path >>
      (fn name => Toplevel.theory (fn state => export-phases state name)))
›
```

ML-file *rewrites.ML*

9.3.1 Semantic Preservation Obligation

fun *rewrite-preservation* :: *IRExpr Rewrite* ⇒ *bool* **where**
rewrite-preservation (*Transform x y*) = ($y \leq x$) |
rewrite-preservation (*Conditional x y cond*) = (*cond* ⟶ ($y \leq x$)) |
rewrite-preservation (*Sequential x y*) = (*rewrite-preservation x* ∧ *rewrite-preservation y*) |
rewrite-preservation (*Transitive x*) = *rewrite-preservation x*

9.3.2 Termination Obligation

fun *rewrite-termination* :: *IRExpr Rewrite* ⇒ (*IRExpr* ⇒ *nat*) ⇒ *bool* **where**
rewrite-termination (*Transform x y*) *trm* = (*trm x* > *trm y*) |
rewrite-termination (*Conditional x y cond*) *trm* = (*cond* ⟶ (*trm x* > *trm y*)) |

rewrite-termination (*Sequential x y*) *trm* = (*rewrite-termination x trm* ∧ *rewrite-termination y trm*) |
rewrite-termination (*Transitive x*) *trm* = *rewrite-termination x trm*

fun *intval* :: *Value Rewrite* ⇒ *bool* **where**
intval (*Transform x y*) = (*x* ≠ *UndefVal* ∧ *y* ≠ *UndefVal* ⟶ *x* = *y*) |
intval (*Conditional x y cond*) = (*cond* ⟶ (*x* = *y*)) |
intval (*Sequential x y*) = (*intval x* ∧ *intval y*) |
intval (*Transitive x*) = *intval x*

9.3.3 Standard Termination Measure

fun *size* :: *IRExpr* ⇒ *nat* **where**
unary-size:
size (*UnaryExpr op x*) = (*size x*) + *2* |

bin-const-size:
size (*BinaryExpr op x* (*ConstantExpr cy*)) = (*size x*) + *2* |
bin-size:
size (*BinaryExpr op x y*) = (*size x*) + (*size y*) + *2* |
cond-size:
size (*ConditionalExpr c t f*) = (*size c*) + (*size t*) + (*size f*) + *2* |
const-size:
size (*ConstantExpr c*) = *1* |
param-size:
size (*ParameterExpr ind s*) = *2* |
leaf-size:
size (*LeafExpr nid s*) = *2* |
size (*ConstantVar c*) = *2* |
size (*VariableExpr x s*) = *2*

9.3.4 Automated Tactics

named-theorems *size-simps size simplication rules*

method *unfold-optimization* =
(*unfold rewrite-preservation.simps*, *unfold rewrite-termination.simps*,
unfold intval.simps,
rule conjE, *simp*, *simp del*: *le-expr-def*, *force?*)
| (*unfold rewrite-preservation.simps*, *unfold rewrite-termination.simps*,
rule conjE, *simp*, *simp del*: *le-expr-def*, *force?*)

method *unfold-size* =
(((*unfold size.simps*, *simp add*: *size-simps del*: *le-expr-def*)?
; (*simp add*: *size-simps del*: *le-expr-def*)?
; (*auto simp*: *size-simps*)?
; (*unfold size.simps*)?)[*1*])

print-methods

ML ‹
structure System : *RewriteSystem* =
struct
val preservation = @{*const rewrite-preservation*};
val termination = @{*const rewrite-termination*};
val intval = @{*const intval*};
end

structure DSL = *DSL-Rewrites*(*System*);

val - =
  Outer-Syntax.local-theory-to-proof ***command-keyword***‹*optimization*›
    define an optimization and open proof obligation
    (*Parse-Spec.thm-name* : −− *Parse.term*
        >> *DSL.rewrite-cmd*);
›

end

10 Canonicalization Optimizations

theory *Common*
  imports
    OptimizationDSL.Canonicalization
    Semantics.IRTreeEvalThms
begin

lemma *size-pos*[*size-simps*]: $0 < size\ y$
  ⟨*proof*⟩

lemma *size-non-add*[*size-simps*]: $size\ (BinaryExpr\ op\ a\ b) = size\ a + size\ b + 2 \longleftrightarrow \neg(is\text{-}ConstantExpr\ b)$
  ⟨*proof*⟩

lemma *size-non-const*[*size-simps*]:
  $\neg\ is\text{-}ConstantExpr\ y \Longrightarrow 1 < size\ y$
  ⟨*proof*⟩

lemma *size-binary-const*[*size-simps*]:
  $size\ (BinaryExpr\ op\ a\ b) = size\ a + 2 \longleftrightarrow (is\text{-}ConstantExpr\ b)$
  ⟨*proof*⟩

lemma *size-flip-binary*[*size-simps*]:
  $\neg(is\text{-}ConstantExpr\ y) \longrightarrow size\ (BinaryExpr\ op\ (ConstantExpr\ x)\ y) > size\ (BinaryExpr\ op\ y\ (ConstantExpr\ x))$
  ⟨*proof*⟩

lemma *size-binary-lhs-a*[*size-simps*]:

size (*BinaryExpr op* (*BinaryExpr op′ a b*) *c*) > *size a*
⟨*proof*⟩

lemma *size-binary-lhs-b*[*size-simps*]:
size (*BinaryExpr op* (*BinaryExpr op′ a b*) *c*) > *size b*
⟨*proof*⟩

lemma *size-binary-lhs-c*[*size-simps*]:
size (*BinaryExpr op* (*BinaryExpr op′ a b*) *c*) > *size c*
⟨*proof*⟩

lemma *size-binary-rhs-a*[*size-simps*]:
size (*BinaryExpr op c* (*BinaryExpr op′ a b*)) > *size a*
⟨*proof*⟩

lemma *size-binary-rhs-b*[*size-simps*]:
size (*BinaryExpr op c* (*BinaryExpr op′ a b*)) > *size b*
⟨*proof*⟩

lemma *size-binary-rhs-c*[*size-simps*]:
size (*BinaryExpr op c* (*BinaryExpr op′ a b*)) > *size c*
⟨*proof*⟩

lemma *size-binary-lhs*[*size-simps*]:
size (*BinaryExpr op x y*) > *size x*
⟨*proof*⟩

lemma *size-binary-rhs*[*size-simps*]:
size (*BinaryExpr op x y*) > *size y*
⟨*proof*⟩

lemmas *arith*[*size-simps*] = *Suc-leI add-strict-increasing order-less-trans trans-less-add2*

definition *well-formed-equal* :: *Value* ⇒ *Value* ⇒ *bool*
(**infix** ≈ *50*) **where**
well-formed-equal v_1 v_2 = (v_1 ≠ *UndefVal* ⟶ v_1 = v_2)

lemma *well-formed-equal-defn* [*simp*]:
well-formed-equal v_1 v_2 = (v_1 ≠ *UndefVal* ⟶ v_1 = v_2)
⟨*proof*⟩

end

10.1 AbsNode Phase

theory *AbsPhase*
imports
Common

begin

phase *AbsNode*
  terminating *size*
begin

lemma *abs-pos*:
  fixes $v :: ('a :: len\ word)$
  assumes $0 \le s\ v$
  shows $(if\ v <s\ 0\ then\ -\ v\ else\ v) = v$
  ⟨*proof*⟩

lemma *abs-neg*:
  fixes $v :: ('a :: len\ word)$
  assumes $v <s\ 0$
  assumes $-(2\ \hat{}\ (Nat.size\ v\ -\ 1)) <s\ v$
  shows $(if\ v <s\ 0\ then\ -\ v\ else\ v) = -\ v \wedge 0 <s\ -v$
  ⟨*proof*⟩

lemma *abs-max-neg*:
  fixes $v :: ('a :: len\ word)$
  assumes $v <s\ 0$
  assumes $-\ (2\ \hat{}\ (Nat.size\ v\ -\ 1)) = v$
  shows $-v = v$
  ⟨*proof*⟩

lemma *final-abs*:
  fixes $v :: ('a :: len\ word)$
  assumes $take\text{-}bit\ (Nat.size\ v)\ v = v$
  assumes $-\ (2\ \hat{}\ (Nat.size\ v\ -\ 1)) \neq v$
  shows $0 \le s\ (if\ v <s\ 0\ then\ -v\ else\ v)$

⟨*proof*⟩

lemma *wf-abs*: $is\text{-}IntVal\ x \Longrightarrow intval\text{-}abs\ x \neq UndefVal$
  ⟨*proof*⟩

fun *bin-abs* :: $'a\ ::len\ word \Rightarrow\ 'a\ ::len\ word$ **where**
  $bin\text{-}abs\ v = (if\ (v <s\ 0)\ then\ (-\ v)\ else\ v)$

lemma *val-abs-zero*:
  $intval\text{-}abs\ (new\text{-}int\ b\ 0) = new\text{-}int\ b\ 0$

⟨*proof*⟩

lemma *less-eq-zero*:
assumes $val\text{-}to\text{-}bool\ (val[(IntVal\ b\ 0) < (IntVal\ b\ v)])$
shows $int\text{-}signed\text{-}value\ b\ v > 0$
⟨*proof*⟩

lemma *val-abs-pos*:
assumes $val\text{-}to\text{-}bool(val[(new\text{-}int\ b\ 0) < (new\text{-}int\ b\ v)])$
shows $intval\text{-}abs\ (new\text{-}int\ b\ v) = (new\text{-}int\ b\ v)$
⟨*proof*⟩

lemma *val-abs-neg*:
assumes $val\text{-}to\text{-}bool(val[(new\text{-}int\ b\ v) < (new\text{-}int\ b\ 0)])$
shows $intval\text{-}abs\ (new\text{-}int\ b\ v) = intval\text{-}negate\ (new\text{-}int\ b\ v)$
⟨*proof*⟩

lemma *val-bool-unwrap*:
$val\text{-}to\text{-}bool\ (bool\text{-}to\text{-}val\ v) = v$
⟨*proof*⟩

lemma *take-bit-unwrap*:
$b = 64 \Longrightarrow take\text{-}bit\ b\ (v1{::}64\ word) = v1$
⟨*proof*⟩

lemma *bit-less-eq-def*:
fixes $v1\ v2 :: 64\ word$
assumes $b \leq 64$
shows $sint\ (signed\text{-}take\text{-}bit\ (b - Suc\ (0{::}nat))\ (take\text{-}bit\ b\ v1))$
$< sint\ (signed\text{-}take\text{-}bit\ (b - Suc\ (0{::}nat))\ (take\text{-}bit\ b\ v2)) \longleftrightarrow$
$signed\text{-}take\text{-}bit\ (63{::}nat)\ (Word.rep\ v1) < signed\text{-}take\text{-}bit\ (63{::}nat)\ (Word.rep$
$v2)$
⟨*proof*⟩

lemma *less-eq-def*:

shows $val\text{-}to\text{-}bool(val[(new\text{-}int\ b\ v1) < (new\text{-}int\ b\ v2)]) \longleftrightarrow v1 <s\ v2$
⟨*proof*⟩

lemma *val-abs-always-pos*:
assumes $intval\text{-}abs\ (new\text{-}int\ b\ v) = (new\text{-}int\ b\ v')$
shows $0 \leq s\ v'$
⟨*proof*⟩

lemma *intval-abs-elims*:
assumes $intval\text{-}abs\ x \neq UndefVal$
shows $\exists\, t\ v\ .\ x = IntVal\ t\ v \wedge intval\text{-}abs\ x = new\text{-}int\ t\ (if\ int\text{-}signed\text{-}value\ t\ v <$
$0\ then\ -\ v\ else\ v)$

⟨*proof*⟩

lemma *wf-abs-new-int*:
assumes *intval-abs* (*IntVal t v*) $\neq$ *UndefVal*
shows *intval-abs* (*IntVal t v*) = *new-int t v* $\vee$ *intval-abs* (*IntVal t v*) = *new-int t* ($-v$)
⟨*proof*⟩

lemma *mono-undef-abs*:
assumes *intval-abs* (*intval-abs x*) $\neq$ *UndefVal*
shows *intval-abs x* $\neq$ *UndefVal*
⟨*proof*⟩

lemma *val-abs-idem*:
assumes *intval-abs*(*intval-abs*(*x*)) $\neq$ *UndefVal*
shows *intval-abs*(*intval-abs*(*x*)) = *intval-abs x*
⟨*proof*⟩

lemma *val-abs-negate*:
assumes *intval-abs* (*intval-negate x*) $\neq$ *UndefVal*
shows *intval-abs* (*intval-negate x*) = *intval-abs x*
⟨*proof*⟩

Optimisations

optimization *AbsIdempotence*: $abs(abs(x)) \longmapsto abs(x)$
⟨*proof*⟩

optimization *AbsNegate*: $(abs(-x)) \longmapsto abs(x)$
⟨*proof*⟩

end

end

10.2 AddNode Phase

theory *AddPhase*
imports
Common
begin

phase *AddNode*
terminating *size*
begin

lemma *binadd-commute*:
assumes *bin-eval BinAdd x y* $\neq$ *UndefVal*
shows *bin-eval BinAdd x y* = *bin-eval BinAdd y x*

⟨*proof*⟩

optimization *AddShiftConstantRight*: ((*const v*) + *y*) $\longmapsto$ *y* + (*const v*) *when* ¬(*is-ConstantExpr y*)
⟨*proof*⟩

optimization *AddShiftConstantRight2*: ((*const v*) + *y*) $\longmapsto$ *y* + (*const v*) *when* ¬(*is-ConstantExpr y*)
⟨*proof*⟩

lemma *is-neutral-0* [*simp*]:
assumes *1*: *intval-add* (*IntVal b x*) (*IntVal b 0*) ≠ *UndefVal*
shows *intval-add* (*IntVal b x*) (*IntVal b 0*) = (*new-int b x*)
⟨*proof*⟩

optimization *AddNeutral*: (*e* + (*const* (*IntVal 32 0*))) $\longmapsto$ *e*
⟨*proof*⟩

ML-val ‹@{*term* ‹*x* = *y*›}›

lemma *NeutralLeftSubVal*:
assumes *e1* = *new-int b ival*
shows *val*[(*e1* − *e2*) + *e2*] ≈ *e1*
⟨*proof*⟩

optimization *RedundantSubAdd*: ((e_1 − e_2) + e_2) $\longmapsto$ e_1
⟨*proof*⟩

lemma *allE2*: (∀ *x y*. *P x y*) ⟹ (*P a b* ⟹ *R*) ⟹ *R*
⟨*proof*⟩

lemma *just-goal2*:
assumes *1*: (∀ *a b*. (*intval-add* (*intval-sub a b*) *b* ≠ *UndefVal* ∧ *a* ≠ *UndefVal* ⟶
intval-add (*intval-sub a b*) *b* = *a*))
shows (*BinaryExpr BinAdd* (*BinaryExpr BinSub* e_1 e_2) e_2) ≥ e_1
⟨*proof*⟩

optimization *RedundantSubAdd2*: $e_2 + (e_1 - e_2) \longmapsto e_1$
⟨*proof*⟩

lemma *AddToSubHelperLowLevel*:
shows *intval-add* (*intval-negate* e) y = *intval-sub* y e (**is** $?x = ?y$)
⟨*proof*⟩

print-phases

lemma *val-redundant-add-sub*:
assumes a = *new-int bb ival*
assumes $val[b + a] \neq UndefVal$
shows $val[(b + a) - b] = a$
⟨*proof*⟩

lemma *val-add-right-negate-to-sub*:
assumes $val[x + e] \neq UndefVal$
shows $val[x + (-e)] = val[x - e]$
⟨*proof*⟩

lemma *exp-add-left-negate-to-sub*:
$exp[-e + y] \geq exp[y - e]$
⟨*proof*⟩

Optimisations

optimization *RedundantAddSub*: $(b + a) - b \longmapsto a$
⟨*proof*⟩

optimization *AddRightNegateToSub*: $x + -e \longmapsto x - e$
⟨*proof*⟩

optimization *AddLeftNegateToSub*: $-e + y \longmapsto y - e$
⟨*proof*⟩

end

end

10.3 AndNode Phase

theory *AndPhase*
  imports
    Common
    Proofs.StampEvalThms
begin

context *stamp-mask*
begin

lemma *AndRightFallthrough*: $(((and\ (not\ (\downarrow x))\ (\uparrow y)) = 0)) \longrightarrow exp[x\ \&\ y] \geq exp[y]$
  ⟨*proof*⟩

lemma *AndLeftFallthrough*: $(((and\ (not\ (\downarrow y))\ (\uparrow x)) = 0)) \longrightarrow exp[x\ \&\ y] \geq exp[x]$
  ⟨*proof*⟩
end

phase *AndNode*
  terminating *size*
begin

lemma *bin-and-nots*:
 $(\sim x\ \&\ \sim y) = (\sim(x \mid y))$
  ⟨*proof*⟩

lemma *bin-and-neutral*:
 $(x\ \&\ \sim False) = x$
  ⟨*proof*⟩

lemma *val-and-equal*:
  assumes $x = new\text{-}int\ b\ v$
  and $val[x\ \&\ x] \neq UndefVal$
  shows $val[x\ \&\ x] = x$
   ⟨*proof*⟩

lemma *val-and-nots*:
  $val[\sim x\ \&\ \sim y] = val[\sim(x \mid y)]$
  ⟨*proof*⟩

lemma *val-and-neutral*:
  assumes $x = \textit{new-int}\ b\ v$
  and $\textit{val}[x\ \&\ {\sim}(\textit{new-int}\ b'\ 0)] \neq \textit{UndefVal}$
  shows $\textit{val}[x\ \&\ {\sim}(\textit{new-int}\ b'\ 0)] = x$
  ⟨*proof*⟩

lemma *val-and-zero*:
  assumes $x = \textit{new-int}\ b\ v$
  shows $\textit{val}[x\ \&\ (\textit{IntVal}\ b\ 0)] = \textit{IntVal}\ b\ 0$
  ⟨*proof*⟩

lemma *exp-and-equal*:
  $\textit{exp}[x\ \&\ x] \geq \textit{exp}[x]$
  ⟨*proof*⟩

lemma *exp-and-nots*:
  $\textit{exp}[{\sim}x\ \&\ {\sim}y] \geq \textit{exp}[{\sim}(x\ |\ y)]$
  ⟨*proof*⟩

lemma *exp-sign-extend*:
  assumes $e = (1 << \textit{In}) - 1$
  shows $\textit{BinaryExpr BinAnd}\ (\textit{UnaryExpr}\ (\textit{UnarySignExtend In Out})\ x)$
  $(\textit{ConstantExpr}\ (\textit{new-int}\ b\ e))$
  $\geq (\textit{UnaryExpr}\ (\textit{UnaryZeroExtend In Out})\ x)$
  ⟨*proof*⟩

lemma *val-and-commute*[*simp*]:
  $\textit{val}[x\ \&\ y] = \textit{val}[y\ \&\ x]$
  ⟨*proof*⟩

Optimisations

optimization *AndEqual*: $x\ \&\ x \longmapsto x$
  ⟨*proof*⟩

optimization *AndShiftConstantRight*: $((\textit{const}\ x)\ \&\ y) \longmapsto y\ \&\ (\textit{const}\ x)$
  $\textit{when}\ \neg(\textit{is-ConstantExpr}\ y)$
  ⟨*proof*⟩

optimization *AndNots*: $(\sim x)$ & $(\sim y) \longmapsto \sim(x \mid y)$
⟨*proof*⟩

optimization *AndSignExtend*: *BinaryExpr BinAnd* (*UnaryExpr* (*UnarySignExtend In Out*) (*x*))
(*const* (*new-int b e*))
$\longmapsto$ (*UnaryExpr* (*UnaryZeroExtend In Out*) (*x*))
when (*e* = (*1* << *In*) − *1*)
⟨*proof*⟩

optimization *AndNeutral*: (*x* & $\sim$(*const* (*IntVal b 0*))) $\longmapsto$ *x*
when (*wf-stamp x* ∧ *stamp-expr x* = *IntegerStamp b lo hi*)
⟨*proof*⟩

optimization *AndRightFallThrough*: (*x* & *y*) $\longmapsto$ *y*
when (((*and* (*not* (*IRExpr-down x*)) (*IRExpr-up y*)) = *0*))
⟨*proof*⟩

optimization *AndLeftFallThrough*: (*x* & *y*) $\longmapsto$ *x*
when (((*and* (*not* (*IRExpr-down y*)) (*IRExpr-up x*)) = *0*))
⟨*proof*⟩

end

end

10.4 BinaryNode Phase

theory *BinaryNode*
imports
Common
begin

phase *BinaryNode*
terminating *size*
begin

optimization *BinaryFoldConstant*: *BinaryExpr op* (*const v1*) (*const v2*) $\longmapsto$ *ConstantExpr* (*bin-eval op v1 v2*)
⟨*proof*⟩

print-facts

end

end

10.5 ConditionalNode Phase

theory *ConditionalPhase*
  imports
    Common
    Proofs.StampEvalThms
begin

phase *ConditionalNode*
  terminating *size*
begin

lemma *negates*: $\exists\, v\ b.\ e = \mathit{IntVal}\ b\ v \wedge b > 0 \Longrightarrow \mathit{val\text{-}to\text{-}bool}\ (\mathit{val}[e]) \longleftrightarrow \neg(\mathit{val\text{-}to\text{-}bool}\ (\mathit{val}[!e]))$
  ⟨*proof*⟩

lemma *negation-condition-intval*:
  assumes $e = \mathit{IntVal}\ b\ ie$
  assumes $0 < b$
  shows $\mathit{val}[(!e)\ ?\ x : y] = \mathit{val}[e\ ?\ y : x]$
  ⟨*proof*⟩

lemma *negation-preserve-eval*:
  assumes $[m, p] \vdash \mathit{exp}[!e] \mapsto v$
  shows $\exists\, v'.\ ([m, p] \vdash \mathit{exp}[e] \mapsto v') \wedge v = \mathit{val}[!v']$
  ⟨*proof*⟩

lemma *negation-preserve-eval-intval*:
  assumes $[m, p] \vdash \mathit{exp}[!e] \mapsto v$
  shows $\exists\, v'\ b\ vv.\ ([m, p] \vdash \mathit{exp}[e] \mapsto v') \wedge v' = \mathit{IntVal}\ b\ vv \wedge b > 0$
  ⟨*proof*⟩

optimization *NegateConditionFlipBranches*: $((!e)\ ?\ x : y) \longmapsto (e\ ?\ y : x)$
  ⟨*proof*⟩

optimization *DefaultTrueBranch*: $(\mathit{true}\ ?\ x : y) \longmapsto x$ ⟨*proof*⟩

optimization *DefaultFalseBranch*: $(\mathit{false}\ ?\ x : y) \longmapsto y$ ⟨*proof*⟩

optimization *ConditionalEqualBranches*: $(e\ ?\ x : x) \longmapsto x$ ⟨*proof*⟩

optimization *condition-bounds-x*: $((u < v)\ ?\ x : y) \longmapsto x$
    when $(\mathit{stamp\text{-}under}\ (\mathit{stamp\text{-}expr}\ u)\ (\mathit{stamp\text{-}expr}\ v) \wedge \mathit{wf\text{-}stamp}\ u \wedge \mathit{wf\text{-}stamp}\ v)$
  ⟨*proof*⟩

optimization *condition-bounds-y*: $((u < v)\ ?\ x : y) \longmapsto y$
    when $(\mathit{stamp\text{-}under}\ (\mathit{stamp\text{-}expr}\ v)\ (\mathit{stamp\text{-}expr}\ u) \wedge \mathit{wf\text{-}stamp}\ u \wedge \mathit{wf\text{-}stamp}\ v)$
  ⟨*proof*⟩

lemma *val-optimise-integer-test*:
assumes $\exists\, v.\ x = IntVal\ 32\ v$
shows $val[((x\ \&\ (IntVal\ 32\ 1))\ eq\ (IntVal\ 32\ 0))\ ?\ (IntVal\ 32\ 0) : (IntVal\ 32\ 1)]$
$=$
$val[x\ \&\ IntVal\ 32\ 1]$
$\langle proof \rangle$

optimization *ConditionalEliminateKnownLess*: $((x < y)\ ?\ x : y) \longmapsto x$
$when\ (stamp\text{-}under\ (stamp\text{-}expr\ x)\ (stamp\text{-}expr\ y)$
$\wedge\ wf\text{-}stamp\ x \wedge wf\text{-}stamp\ y)$
$\langle proof \rangle$

optimization *ConditionalEqualIsRHS*: $((x\ eq\ y)\ ?\ x : y) \longmapsto y$
$\langle proof \rangle$

optimization *normalizeX*: $((x\ eq\ const\ (IntVal\ 32\ 0))\ ?$
$(const\ (IntVal\ 32\ 0)) : (const\ (IntVal\ 32\ 1))) \longmapsto x$
$when\ (IRExpr\text{-}up\ x = 1) \wedge stamp\text{-}expr\ x = IntegerStamp$
$b\ 0\ 1$
$\langle proof \rangle$

optimization *normalizeX2*: $((x\ eq\ (const\ (IntVal\ 32\ 1)))\ ?$
$(const\ (IntVal\ 32\ 1)) : (const\ (IntVal\ 32\ 0))) \longmapsto x$
$when\ (x = ConstantExpr\ (IntVal\ 32\ 0) \mid (x = ConstantExpr$
$(IntVal\ 32\ 1)))$ $\langle proof \rangle$

optimization *flipX*: $((x\ eq\ (const\ (IntVal\ 32\ 0)))\ ?$
$(const\ (IntVal\ 32\ 1)) : (const\ (IntVal\ 32\ 0))) \longmapsto$
$x \oplus (const\ (IntVal\ 32\ 1))$
$when\ (x = ConstantExpr\ (IntVal\ 32\ 0) \mid (x = ConstantExpr$
$(IntVal\ 32\ 1)))$ $\langle proof \rangle$

optimization *flipX2*: $((x\ eq\ (const\ (IntVal\ 32\ 1)))\ ?$
$(const\ (IntVal\ 32\ 0)) : (const\ (IntVal\ 32\ 1))) \longmapsto$
$x \oplus (const\ (IntVal\ 32\ 1))$
$when\ (x = ConstantExpr\ (IntVal\ 32\ 0) \mid (x = ConstantExpr$
$(IntVal\ 32\ 1)))$ $\langle proof \rangle$

lemma *stamp-of-default*:
assumes $stamp\text{-}expr\ x = default\text{-}stamp$
assumes $wf\text{-}stamp\ x$
shows $([m, p] \vdash x \mapsto v) \longrightarrow (\exists\, vv.\ v = IntVal\ 32\ vv)$

⟨*proof*⟩

optimization *OptimiseIntegerTest*:
$(((x \mathbin{\&} (const\ (IntVal\ 32\ 1)))\ eq\ (const\ (IntVal\ 32\ 0)))\ ?$
$(const\ (IntVal\ 32\ 0)) : (const\ (IntVal\ 32\ 1))) \longmapsto$
$x \mathbin{\&} (const\ (IntVal\ 32\ 1))$
$when\ (stamp\text{-}expr\ x = default\text{-}stamp \wedge wf\text{-}stamp\ x)$
⟨*proof*⟩

optimization *opt-optimise-integer-test-2*:
$(((x \mathbin{\&} (const\ (IntVal\ 32\ 1)))\ eq\ (const\ (IntVal\ 32\ 0)))\ ?$
$(const\ (IntVal\ 32\ 0)) : (const\ (IntVal\ 32\ 1))) \longmapsto$
x
$when\ (x = ConstantExpr\ (IntVal\ 32\ 0) \mid (x = ConstantExpr\ (IntVal\ 32\ 1)))$ ⟨*proof*⟩

end

end

10.6 MulNode Phase

theory *MulPhase*
imports
Common
Proofs.StampEvalThms
begin

fun *mul-size* :: *IRExpr* ⇒ *nat* **where**
$mul\text{-}size\ (UnaryExpr\ op\ e) = (mul\text{-}size\ e) + 2 \mid$
$mul\text{-}size\ (BinaryExpr\ BinMul\ x\ y) = ((mul\text{-}size\ x) + (mul\text{-}size\ y) + 2) * 2 \mid$
$mul\text{-}size\ (BinaryExpr\ op\ x\ y) = (mul\text{-}size\ x) + (mul\text{-}size\ y) + 2 \mid$
$mul\text{-}size\ (ConditionalExpr\ cond\ t\ f) = (mul\text{-}size\ cond) + (mul\text{-}size\ t) + (mul\text{-}size\ f) + 2 \mid$
$mul\text{-}size\ (ConstantExpr\ c) = 1 \mid$
$mul\text{-}size\ (ParameterExpr\ ind\ s) = 2 \mid$
$mul\text{-}size\ (LeafExpr\ nid\ s) = 2 \mid$
$mul\text{-}size\ (ConstantVar\ c) = 2 \mid$
$mul\text{-}size\ (VariableExpr\ x\ s) = 2$

phase *MulNode*
terminating *mul-size*

begin

lemma *bin-eliminate-redundant-negative*:
$uminus\ (x :: 'a::len\ word) * uminus\ (y :: 'a::len\ word) = x * y$
⟨*proof*⟩

lemma *bin-multiply-identity*:
$(x :: 'a::len\ word) * 1 = x$
⟨*proof*⟩

lemma *bin-multiply-eliminate*:
$(x :: 'a::len\ word) * 0 = 0$
⟨*proof*⟩

lemma *bin-multiply-negative*:
$(x :: 'a::len\ word) * uminus\ 1 = uminus\ x$
⟨*proof*⟩

lemma *bin-multiply-power-2*:
$(x:: 'a::len\ word) * (2\hat{}j) = x << j$
⟨*proof*⟩

lemma *take-bit64*[*simp*]:
fixes $w :: int64$
shows $take\text{-}bit\ 64\ w = w$
⟨*proof*⟩

lemma *mergeTakeBit*:
fixes $a :: nat$
fixes $b\ c :: 64\ word$
shows $take\text{-}bit\ a\ (take\text{-}bit\ a\ (b) * take\text{-}bit\ a\ (c)) =$
$take\text{-}bit\ a\ (b * c)$
⟨*proof*⟩

lemma *val-eliminate-redundant-negative*:
assumes $val[-x * -y] \neq UndefVal$
shows $val[-x * -y] = val[x * y]$
⟨*proof*⟩

lemma *val-multiply-neutral*:
assumes $x = new\text{-}int\ b\ v$
shows $val[x * (IntVal\ b\ 1)] = val[x]$
⟨*proof*⟩

lemma *val-multiply-zero*:
  assumes $x = \textit{new-int}\ b\ v$
  shows $\textit{val}[x * (\textit{IntVal}\ b\ 0)] = \textit{IntVal}\ b\ 0$
  ⟨*proof*⟩

lemma *val-multiply-negative*:
  assumes $x = \textit{new-int}\ b\ v$
  shows $\textit{val}[x * \textit{intval-negate}\ (\textit{IntVal}\ b\ 1)] = \textit{intval-negate}\ x$
  ⟨*proof*⟩

lemma *val-MulPower2*:
  fixes $i :: 64\ \textit{word}$
  assumes $y = \textit{IntVal}\ 64\ (2\ \hat{}\ \textit{unat}(i))$
  and $0 < i$
  and $i < 64$
  and $\textit{val}[x * y] \neq \textit{UndefVal}$
  shows $\textit{val}[x * y] = \textit{val}[x << \textit{IntVal}\ 64\ i]$
  ⟨*proof*⟩

lemma *val-MulPower2Add1*:
  fixes $i :: 64\ \textit{word}$
  assumes $y = \textit{IntVal}\ 64\ ((2\ \hat{}\ \textit{unat}(i)) + 1)$
  and $0 < i$
  and $i < 64$
  and $\textit{val-to-bool}(\textit{val}[\textit{IntVal}\ 64\ 0 < x])$
  and $\textit{val-to-bool}(\textit{val}[\textit{IntVal}\ 64\ 0 < y])$
  shows $\textit{val}[x * y] = \textit{val}[(x << \textit{IntVal}\ 64\ i) + x]$
  ⟨*proof*⟩

lemma *val-MulPower2Sub1*:
  fixes $i :: 64\ \textit{word}$
  assumes $y = \textit{IntVal}\ 64\ ((2\ \hat{}\ \textit{unat}(i)) - 1)$
  and $0 < i$
  and $i < 64$
  and $\textit{val-to-bool}(\textit{val}[\textit{IntVal}\ 64\ 0 < x])$
  and $\textit{val-to-bool}(\textit{val}[\textit{IntVal}\ 64\ 0 < y])$
  shows $\textit{val}[x * y] = \textit{val}[(x << \textit{IntVal}\ 64\ i) - x]$
  ⟨*proof*⟩

lemma *val-distribute-multiplication*:
  assumes $x = \textit{new-int}\ 64\ xx \wedge q = \textit{new-int}\ 64\ qq \wedge a = \textit{new-int}\ 64\ aa$
  shows $\textit{val}[x * (q + a)] = \textit{val}[(x * q) + (x * a)]$
  ⟨*proof*⟩

lemma *val-MulPower2AddPower2*:
fixes $i\ j$ :: *64 word*
assumes $y = IntVal\ 64\ ((2\ \widehat{}\ unat(i)) + (2\ \widehat{}\ unat(j)))$
and $0 < i$
and $0 < j$
and $i < 64$
and $j < 64$
and $x = new\text{-}int\ 64\ xx$
shows $val[x * y] = val[(x << IntVal\ 64\ i) + (x << IntVal\ 64\ j)]$
⟨*proof*⟩

thm-oracles *val-MulPower2AddPower2*

lemma *exp-multiply-zero-64*:
$exp[x * (const\ (IntVal\ 64\ 0))] \geq ConstantExpr\ (IntVal\ 64\ 0)$
⟨*proof*⟩

lemma *exp-multiply-neutral*:
$exp[x * (const\ (IntVal\ b\ 1))] \geq x$
⟨*proof*⟩

thm-oracles *exp-multiply-neutral*

lemma *exp-MulPower2*:
fixes i :: *64 word*
assumes $y = ConstantExpr\ (IntVal\ 64\ (2\ \widehat{}\ unat(i)))$
and $0 < i$
and $i < 64$
and $exp[x > (const\ IntVal\ b\ 0)]$
and $exp[y > (const\ IntVal\ b\ 0)]$
shows $exp[x * y] \geq exp[x << ConstantExpr\ (IntVal\ 64\ i)]$
⟨*proof*⟩

lemma *exp-MulPower2Add1*:
fixes i :: *64 word*
assumes $y = ConstantExpr\ (IntVal\ 64\ ((2\ \widehat{}\ unat(i)) + 1))$
and $0 < i$
and $i < 64$
and $exp[x > (const\ IntVal\ b\ 0)]$
and $exp[y > (const\ IntVal\ b\ 0)]$
shows $exp[x * y] \geq exp[(x << ConstantExpr\ (IntVal\ 64\ i)) + x]$
⟨*proof*⟩

lemma *exp-MulPower2Sub1*:
fixes i :: *64 word*
assumes $y = ConstantExpr\ (IntVal\ 64\ ((2\ \widehat{}\ unat(i)) - 1))$

and $0 < i$
and $i < 64$
and $exp[x > (const\ IntVal\ b\ 0)]$
and $exp[y > (const\ IntVal\ b\ 0)]$
shows $exp[x * y] \geq exp[(x << ConstantExpr\ (IntVal\ 64\ i)) - x]$
⟨*proof*⟩

lemma *exp-MulPower2AddPower2*:
fixes $i\ j :: 64\ word$
assumes $y = ConstantExpr\ (IntVal\ 64\ ((2\ \hat{}\ unat(i)) + (2\ \hat{}\ unat(j))))$
and $0 < i$
and $0 < j$
and $i < 64$
and $j < 64$
and $exp[x > (const\ IntVal\ b\ 0)]$
and $exp[y > (const\ IntVal\ b\ 0)]$
shows $exp[x * y] \geq exp[(x << ConstantExpr\ (IntVal\ 64\ i)) + (x << ConstantExpr\ (IntVal\ 64\ j))]$
⟨*proof*⟩

lemma *greaterConstant*:
fixes $a\ b :: 64\ word$
assumes $a > b$
and $y = ConstantExpr\ (IntVal\ 64\ a)$
and $x = ConstantExpr\ (IntVal\ 64\ b)$
shows $exp[y > x]$
⟨*proof*⟩

lemma *exp-distribute-multiplication*:
shows $exp[(x * q) + (x * a)] \geq exp[x * (q + a)]$
⟨*proof*⟩

Optimisations

optimization *EliminateRedundantNegative*: $-x * -y \longmapsto x * y$
⟨*proof*⟩

optimization *MulNeutral*: $x * ConstantExpr\ (IntVal\ b\ 1) \longmapsto x$
⟨*proof*⟩

optimization *MulEliminator*: $x * ConstantExpr\ (IntVal\ b\ 0) \longmapsto const\ (IntVal\ b\ 0)$
⟨*proof*⟩

optimization *MulNegate*: $x * -(const\ (IntVal\ b\ 1)) \longmapsto -x$
⟨*proof*⟩

fun *isNonZero* :: *Stamp* ⇒ *bool* **where**
isNonZero (*IntegerStamp b lo hi*) = (*lo* > *0*) |
isNonZero - = *False*

lemma *isNonZero-defn*:
assumes *isNonZero* (*stamp-expr x*)
assumes *wf-stamp x*
shows $([m, p] \vdash x \mapsto v) \longrightarrow (\exists\, vv\ b.\ (v = IntVal\ b\ vv \wedge val\text{-}to\text{-}bool\ val[(IntVal\ b\ 0) < v]))$
⟨*proof*⟩

optimization *MulPower2*: $x * y \longmapsto x << const\ (IntVal\ 64\ i)$
when $(i > 0 \wedge$
$64 > i \wedge$
$y = exp[const\ (IntVal\ 64\ (2\ \hat{}\ unat(i)))])$
⟨*proof*⟩

optimization *MulPower2Add1*: $x * y \longmapsto (x << const\ (IntVal\ 64\ i)) + x$
when $(i > 0 \wedge$
$64 > i \wedge$
$y = ConstantExpr\ (IntVal\ 64\ ((2\ \hat{}\ unat(i)) + 1))\ )$
⟨*proof*⟩

optimization *MulPower2Sub1*: $x * y \longmapsto (x << const\ (IntVal\ 64\ i)) - x$
when $(i > 0 \wedge$
$64 > i \wedge$
$y = ConstantExpr\ (IntVal\ 64\ ((2\ \hat{}\ unat(i)) - 1))\ )$
⟨*proof*⟩

end

end

10.7 Experimental AndNode Phase

theory *NewAnd*
imports
Common
Graph.Long
begin

lemma *bin-distribute-and-over-or*:
$bin[z\ \&\ (x\ |\ y)] = bin[(z\ \&\ x)\ |\ (z\ \&\ y)]$
⟨*proof*⟩

lemma *intval-distribute-and-over-or*:

$val[z \ \& \ (x \mid y)] = val[(z \ \& \ x) \mid (z \ \& \ y)]$
⟨*proof*⟩

lemma *exp-distribute-and-over-or*:
$exp[z \ \& \ (x \mid y)] \geq exp[(z \ \& \ x) \mid (z \ \& \ y)]$
⟨*proof*⟩

lemma *intval-and-commute*:
$val[x \ \& \ y] = val[y \ \& \ x]$
⟨*proof*⟩

lemma *intval-or-commute*:
$val[x \mid y] = val[y \mid x]$
⟨*proof*⟩

lemma *intval-xor-commute*:
$val[x \oplus y] = val[y \oplus x]$
⟨*proof*⟩

lemma *exp-and-commute*:
$exp[x \ \& \ z] \geq exp[z \ \& \ x]$
⟨*proof*⟩

lemma *exp-or-commute*:
$exp[x \mid y] \geq exp[y \mid x]$
⟨*proof*⟩

lemma *exp-xor-commute*:
$exp[x \oplus y] \geq exp[y \oplus x]$
⟨*proof*⟩

lemma *bin-eliminate-y*:
assumes $bin[y \ \& \ z] = 0$
shows $bin[(x \mid y) \ \& \ z] = bin[x \ \& \ z]$
⟨*proof*⟩

lemma *intval-eliminate-y*:
assumes $val[y \ \& \ z] = IntVal \ b \ 0$
shows $val[(x \mid y) \ \& \ z] = val[x \ \& \ z]$
⟨*proof*⟩

lemma *intval-and-associative*:
$val[(x \ \& \ y) \ \& \ z] = val[x \ \& \ (y \ \& \ z)]$
⟨*proof*⟩

lemma *intval-or-associative*:
$val[(x \mid y) \mid z] = val[x \mid (y \mid z)]$

$\langle proof \rangle$

lemma *intval-xor-associative*:
$val[(x \oplus y) \oplus z] = val[x \oplus (y \oplus z)]$
$\langle proof \rangle$

lemma *exp-and-associative*:
$exp[(x \ \& \ y) \ \& \ z] \geq exp[x \ \& \ (y \ \& \ z)]$
$\langle proof \rangle$

lemma *exp-or-associative*:
$exp[(x \mid y) \mid z] \geq exp[x \mid (y \mid z)]$
$\langle proof \rangle$

lemma *exp-xor-associative*:
$exp[(x \oplus y) \oplus z] \geq exp[x \oplus (y \oplus z)]$
$\langle proof \rangle$

lemma *intval-and-absorb-or*:
assumes $\exists\, b\ v\ .\ x = \textit{new-int}\ b\ v$
assumes $val[x \ \& \ (x \mid y)] \neq \textit{UndefVal}$
shows $val[x \ \& \ (x \mid y)] = val[x]$
$\langle proof \rangle$

lemma *intval-or-absorb-and*:
assumes $\exists\, b\ v\ .\ x = \textit{new-int}\ b\ v$
assumes $val[x \mid (x \ \& \ y)] \neq \textit{UndefVal}$
shows $val[x \mid (x \ \& \ y)] = val[x]$
$\langle proof \rangle$

lemma *exp-and-absorb-or*:
$exp[x \ \& \ (x \mid y)] \geq exp[x]$
$\langle proof \rangle$

lemma *exp-or-absorb-and*:
$exp[x \mid (x \ \& \ y)] \geq exp[x]$
$\langle proof \rangle$

lemma
assumes $y = 0$
shows $x + y = or\ x\ y$
$\langle proof \rangle$

lemma *no-overlap-or*:
assumes $and\ x\ y = 0$
shows $x + y = or\ x\ y$

$\langle proof \rangle$

context *stamp-mask*
begin

lemma *intval-up-and-zero-implies-zero*:
assumes $and\ (\uparrow x)\ (\uparrow y) = 0$
assumes $[m, p] \vdash x \mapsto xv$
assumes $[m, p] \vdash y \mapsto yv$
assumes $val[xv\ \&\ yv] \neq UndefVal$
shows $\exists\ b\ .\ val[xv\ \&\ yv] = new\text{-}int\ b\ 0$
$\langle proof \rangle$

lemma *exp-eliminate-y*:
$and\ (\uparrow y)\ (\uparrow z) = 0 \longrightarrow BinaryExpr\ BinAnd\ (BinaryExpr\ BinOr\ x\ y)\ z \geq BinaryExpr\ BinAnd\ x\ z$
$\langle proof \rangle$

lemma *leadingZeroBounds*:
fixes $x :: 'a::len\ word$
assumes $n = numberOfLeadingZeros\ x$
shows $0 \leq n \wedge n \leq Nat.size\ x$
$\langle proof \rangle$

lemma *above-nth-not-set*:
fixes $x :: int64$
assumes $n = 64 - numberOfLeadingZeros\ x$
shows $j > n \longrightarrow \neg(bit\ x\ j)$
$\langle proof \rangle$

no-notation *LogicNegationNotation* (!-)

lemma *zero-horner*:
$horner\text{-}sum\ of\text{-}bool\ 2\ (map\ (\lambda x.\ False)\ xs) = 0$
$\langle proof \rangle$

lemma *zero-map*:
assumes $j \leq n$
assumes $\forall\, i.\ j \leq i \longrightarrow \neg(f\ i)$
shows $map\ f\ [0..<n] = map\ f\ [0..<j]\ @\ map\ (\lambda x.\ False)\ [j..<n]$
$\langle proof \rangle$

lemma *map-join-horner*:
  assumes *map f* [*0..*<*n*] = *map f* [*0..*<*j*] @ *map* (*λx. False*) [*j..*<*n*]
  shows *horner-sum of-bool* (*2::'a::len word*) (*map f* [*0..*<*n*]) = *horner-sum of-bool 2* (*map f* [*0..*<*j*])
⟨*proof*⟩

lemma *split-horner*:
  assumes $j \leq n$
  assumes $\forall\, i.\ j \leq i \longrightarrow \neg(f\ i)$
  shows *horner-sum of-bool* (*2::'a::len word*) (*map f* [*0..*<*n*]) = *horner-sum of-bool 2* (*map f* [*0..*<*j*])
  ⟨*proof*⟩

lemma *transfer-map*:
  assumes $\forall\, i.\ i < n \longrightarrow f\ i = f'\ i$
  shows (*map f* [*0..*<*n*]) = (*map f'* [*0..*<*n*])
  ⟨*proof*⟩

lemma *transfer-horner*:
  assumes $\forall\, i.\ i < n \longrightarrow f\ i = f'\ i$
  shows *horner-sum of-bool* (*2::'a::len word*) (*map f* [*0..*<*n*]) = *horner-sum of-bool 2* (*map f'* [*0..*<*n*])
  ⟨*proof*⟩

lemma *L1*:
  assumes $n = 64 - numberOfLeadingZeros\ (\uparrow z)$
  assumes $[m,\ p] \vdash z \mapsto IntVal\ b\ zv$
  shows $and\ v\ zv = and\ (v\ mod\ 2\widehat{\ }n)\ zv$
⟨*proof*⟩

lemma *up-mask-upper-bound*:
  assumes $[m,\ p] \vdash x \mapsto IntVal\ b\ xv$
  shows $xv \leq (\uparrow x)$
  ⟨*proof*⟩

lemma *L2*:
  assumes $numberOfLeadingZeros\ (\uparrow z) + numberOfTrailingZeros\ (\uparrow y) \geq 64$
  assumes $n = 64 - numberOfLeadingZeros\ (\uparrow z)$
  assumes $[m,\ p] \vdash z \mapsto IntVal\ b\ zv$
  assumes $[m,\ p] \vdash y \mapsto IntVal\ b\ yv$
  shows $yv\ mod\ 2\widehat{\ }n = 0$
⟨*proof*⟩

thm-oracles *L1 L2*

lemma *unfold-binary-width-add*:
  shows $([m,p] \vdash BinaryExpr\ BinAdd\ xe\ ye \mapsto IntVal\ b\ val) = (\exists\ x\ y.$
    $(([m,p] \vdash xe \mapsto IntVal\ b\ x) \wedge$
     $([m,p] \vdash ye \mapsto IntVal\ b\ y) \wedge$

$(IntVal\ b\ val = bin\text{-}eval\ BinAdd\ (IntVal\ b\ x)\ (IntVal\ b\ y)) \wedge$
$(IntVal\ b\ val \neq UndefVal)$
)) (**is** $?L = ?R$)
⟨*proof*⟩

lemma *unfold-binary-width-and*:
shows $([m,p] \vdash BinaryExpr\ BinAnd\ xe\ ye \mapsto IntVal\ b\ val) = (\exists\ x\ y.$
$(([m,p] \vdash xe \mapsto IntVal\ b\ x) \wedge$
$([m,p] \vdash ye \mapsto IntVal\ b\ y) \wedge$
$(IntVal\ b\ val = bin\text{-}eval\ BinAnd\ (IntVal\ b\ x)\ (IntVal\ b\ y)) \wedge$
$(IntVal\ b\ val \neq UndefVal)$
)) (**is** $?L = ?R$)
⟨*proof*⟩

lemma *mod-dist-over-add-right*:
fixes $a\ b\ c :: int64$
fixes $n :: nat$
assumes *1*: $0 < n$
assumes *2*: $n < 64$
shows $(a + b\ mod\ 2\hat{}n)\ mod\ 2\hat{}n = (a + b)\ mod\ 2\hat{}n$
⟨*proof*⟩

lemma *numberOfLeadingZeros-range*:
$0 \leq numberOfLeadingZeros\ n \wedge numberOfLeadingZeros\ n \leq Nat.size\ n$
⟨*proof*⟩

lemma *improved-opt*:
assumes $numberOfLeadingZeros\ (\uparrow z) + numberOfTrailingZeros\ (\uparrow y) \geq 64$
shows $exp[(x + y)\ \&\ z] \geq exp[x\ \&\ z]$
⟨*proof*⟩

thm-oracles *improved-opt*

end

phase *NewAnd*
terminating *size*
begin

optimization *redundant-lhs-y-or*: $((x \mid y)\ \&\ z) \longmapsto x\ \&\ z$
when $(((and\ (IRExpr\text{-}up\ y)\ (IRExpr\text{-}up\ z)) = 0))$
⟨*proof*⟩

optimization *redundant-lhs-x-or*: $((x \mid y)\ \&\ z) \longmapsto y\ \&\ z$

when $(((and\ (IRExpr\text{-}up\ x)\ (IRExpr\text{-}up\ z)) = 0))$
⟨*proof*⟩

optimization *redundant-rhs-y-or*: $(z\ \&\ (x\ |\ y)) \longmapsto z\ \&\ x$
when $(((and\ (IRExpr\text{-}up\ y)\ (IRExpr\text{-}up\ z)) = 0))$
⟨*proof*⟩

optimization *redundant-rhs-x-or*: $(z\ \&\ (x\ |\ y)) \longmapsto z\ \&\ y$
when $(((and\ (IRExpr\text{-}up\ x)\ (IRExpr\text{-}up\ z)) = 0))$
⟨*proof*⟩

end

end

10.8 NotNode Phase

theory *NotPhase*
imports
Common
begin

phase *NotNode*
terminating *size*
begin

lemma *bin-not-cancel*:
$bin[\neg(\neg(e))] = bin[e]$
⟨*proof*⟩

lemma *val-not-cancel*:
assumes $val[\sim(new\text{-}int\ b\ v)] \neq UndefVal$
shows $val[\sim(\sim(new\text{-}int\ b\ v))] = (new\text{-}int\ b\ v)$
⟨*proof*⟩

lemma *exp-not-cancel*:
$exp[\sim(\sim a)] \geq exp[a]$
⟨*proof*⟩

Optimisations

optimization *NotCancel*: $exp[\sim(\sim a)] \longmapsto a$
⟨*proof*⟩

end

end

10.9 OrNode Phase

theory *OrPhase*
  imports
    Common
begin

context *stamp-mask*
begin

Taking advantage of the truth table of or operations.

#	x	y	$x \| y$
1	0	0	0
2	0	1	1
3	1	0	1
4	1	1	1

If row 2 never applies, that is, canBeZero x & canBeOne y = 0, then $(x|y) = x$.

Likewise, if row 3 never applies, canBeZero y & canBeOne x = 0, then $(x|y) = y$.

lemma *OrLeftFallthrough*:
  assumes $(and\ (not\ (\downarrow x))\ (\uparrow y)) = 0$
  shows $exp[x \mid y] \geq exp[x]$
  ⟨*proof*⟩

lemma *OrRightFallthrough*:
  assumes $(and\ (not\ (\downarrow y))\ (\uparrow x)) = 0$
  shows $exp[x \mid y] \geq exp[y]$
  ⟨*proof*⟩

end

phase *OrNode*
  terminating *size*
begin

lemma *bin-or-equal*:
  $bin[x \mid x] = bin[x]$
  ⟨*proof*⟩

lemma *bin-shift-const-right-helper*:
  $x \mid y = y \mid x$

$\langle proof \rangle$

lemma *bin-or-not-operands*:
$(\sim x \mid \sim y) = (\sim(x \ \& \ y))$
$\langle proof \rangle$

lemma *val-or-equal*:
assumes $x = new\text{-}int\ b\ v$
and $(val[x \mid x] \neq UndefVal)$
shows $val[x \mid x] = val[x]$
$\langle proof \rangle$

lemma *val-elim-redundant-false*:
assumes $x = new\text{-}int\ b\ v$
and $val[x \mid false] \neq UndefVal$
shows $val[x \mid false] = val[x]$
$\langle proof \rangle$

lemma *val-shift-const-right-helper*:
$val[x \mid y] = val[y \mid x]$
$\langle proof \rangle$

lemma *val-or-not-operands*:
$val[\sim x \mid \sim y] = val[\sim(x \ \& \ y)]$
$\langle proof \rangle$

lemma *exp-or-equal*:
$exp[x \mid x] \geq exp[x]$
$\langle proof \rangle$

lemma *exp-elim-redundant-false*:
$exp[x \mid false] \geq exp[x]$
$\langle proof \rangle$

Optimisations

optimization *OrEqual*: $x \mid x \longmapsto x$
$\langle proof \rangle$

optimization *OrShiftConstantRight*: $((const\ x) \mid y) \longmapsto y \mid (const\ x)$ *when* $\neg(is\text{-}ConstantExpr\ y)$
$\langle proof \rangle$

optimization *EliminateRedundantFalse*: $x \mid false \longmapsto x$
$\langle proof \rangle$

optimization *OrNotOperands*: $(\sim x \mid \sim y) \longmapsto \sim(x \ \& \ y)$
$\langle proof \rangle$

optimization *OrLeftFallthrough*:
x | *y* ⟼ *x* *when* ((*and* (*not* (*IRExpr-down* *x*)) (*IRExpr-up* *y*)) = *0*)
⟨*proof*⟩

optimization *OrRightFallthrough*:
x | *y* ⟼ *y* *when* ((*and* (*not* (*IRExpr-down* *y*)) (*IRExpr-up* *x*)) = *0*)
⟨*proof*⟩

end

end

10.10 ShiftNode Phase

theory *ShiftPhase*
imports
Common
begin

phase *ShiftNode*
terminating *size*
begin

fun *intval-log2* :: *Value* ⇒ *Value* **where**
intval-log2 (*IntVal* *b* *v*) = *IntVal* *b* (*word-of-int* (*SOME* *e*. *v*=*2*^*e*)) |
intval-log2 - = *UndefVal*

fun *in-bounds* :: *Value* ⇒ *int* ⇒ *int* ⇒ *bool* **where**
in-bounds (*IntVal* *b* *v*) *l* *h* = (*l* < *sint* *v* ∧ *sint* *v* < *h*) |
in-bounds - *l* *h* = *False*

lemma
assumes *in-bounds* (*intval-log2* *val-c*) *0* *32*
shows *intval-left-shift* *x* (*intval-log2* *val-c*) = *intval-mul* *x* *val-c*
⟨*proof*⟩

lemma *e-intval*:
n = *intval-log2* *val-c* ∧ *in-bounds* *n* *0* *32* ⟶
intval-left-shift *x* (*intval-log2* *val-c*) =
intval-mul *x* *val-c*
⟨*proof*⟩

optimization *e*:
x ∗ (*const* *c*) ⟼ *x* << (*const* *n*) *when* (*n* = *intval-log2* *c* ∧ *in-bounds* *n* *0* *32*)
⟨*proof*⟩

end

end

10.11 SignedDivNode Phase

theory *SignedDivPhase*
  imports
    Common
begin

phase *SignedDivNode*
  terminating *size*
begin

lemma *val-division-by-one-is-self-32*:
  assumes $x = \textit{new-int 32 v}$
  shows $\textit{intval-div } x \; (\textit{IntVal 32 1}) = x$
  ⟨*proof*⟩

end

end

10.12 SignedRemNode Phase

theory *SignedRemPhase*
  imports
    Common
begin

phase *SignedRemNode*
  terminating *size*
begin

lemma *val-remainder-one*:
  assumes $\textit{intval-mod } x \; (\textit{IntVal 32 1}) \neq \textit{UndefVal}$
  shows $\textit{intval-mod } x \; (\textit{IntVal 32 1}) = \textit{IntVal 32 0}$
  ⟨*proof*⟩

value $\textit{word-of-int} \; (\textit{sint} \; (\textit{x2::32 word}) \; \textit{smod } 1)$

end

end

10.13 SubNode Phase

theory *SubPhase*
  imports
    Common
    Proofs.StampEvalThms
begin

phase *SubNode*
  terminating *size*
begin

lemma *bin-sub-after-right-add*:
  shows $((x::('a::len)\ word) + (y::('a::len)\ word)) - y = x$
  ⟨*proof*⟩

lemma *sub-self-is-zero*:
  shows $(x::('a::len)\ word) - x = 0$
  ⟨*proof*⟩

lemma *bin-sub-then-left-add*:
  shows $(x::('a::len)\ word) - (x + (y::('a::len)\ word)) = -y$
  ⟨*proof*⟩

lemma *bin-sub-then-left-sub*:
  shows $(x::('a::len)\ word) - (x - (y::('a::len)\ word)) = y$
  ⟨*proof*⟩

lemma *bin-subtract-zero*:
  shows $(x :: 'a::len\ word) - (0 :: 'a::len\ word) = x$
  ⟨*proof*⟩

lemma *bin-sub-negative-value*:
 $(x :: ('a::len)\ word) - (-(y :: ('a::len)\ word)) = x + y$
  ⟨*proof*⟩

lemma *bin-sub-self-is-zero*:
 $(x :: ('a::len)\ word) - x = 0$
  ⟨*proof*⟩

lemma *bin-sub-negative-const*:
$(x :: 'a::len\ word) - (-(y :: 'a::len\ word)) = x + y$
  ⟨*proof*⟩

lemma *val-sub-after-right-add-2*:

assumes $x = new\text{-}int\ b\ v$
assumes $val[(x + y) - y] \neq UndefVal$
shows $val[(x + y) - y] = val[x]$
⟨*proof*⟩

lemma *val-sub-after-left-sub*:
assumes $val[(x - y) - x] \neq UndefVal$
shows $val[(x - y) - x] = val[-y]$
⟨*proof*⟩

lemma *val-sub-then-left-sub*:
assumes $y = new\text{-}int\ b\ v$
assumes $val[x - (x - y)] \neq UndefVal$
shows $val[x - (x - y)] = val[y]$
⟨*proof*⟩

lemma *val-subtract-zero*:
assumes $x = new\text{-}int\ b\ v$
assumes $intval\text{-}sub\ x\ (IntVal\ b\ 0) \neq UndefVal$
shows $intval\text{-}sub\ x\ (IntVal\ b\ 0) = val[x]$
⟨*proof*⟩

lemma *val-zero-subtract-value*:
assumes $x = new\text{-}int\ b\ v$
assumes $intval\text{-}sub\ (IntVal\ b\ 0)\ x \neq UndefVal$
shows $intval\text{-}sub\ (IntVal\ b\ 0)\ x = val[-x]$
⟨*proof*⟩

lemma *val-sub-then-left-add*:
assumes $val[x - (x + y)] \neq UndefVal$
shows $val[x - (x + y)] = val[-y]$
⟨*proof*⟩

lemma *val-sub-negative-value*:
assumes $val[x - (-y)] \neq UndefVal$
shows $val[x - (-y)] = val[x + y]$
⟨*proof*⟩

lemma *val-sub-self-is-zero*:
assumes $x = new\text{-}int\ b\ v \wedge val[x - x] \neq UndefVal$
shows $val[x - x] = new\text{-}int\ b\ 0$
⟨*proof*⟩

lemma *val-sub-negative-const*:
assumes $y = new\text{-}int\ b\ v \wedge val[x - (-y)] \neq UndefVal$
shows $val[x - (-y)] = val[x + y]$
⟨*proof*⟩

lemma *exp-sub-after-right-add*:
shows $exp[(x + y) - y] \geq exp[x]$
⟨*proof*⟩

lemma *exp-sub-after-right-add2*:
shows $exp[(x + y) - x] \geq exp[y]$
⟨*proof*⟩

lemma *exp-sub-negative-value*:
$exp[x - (-y)] \geq exp[x + y]$
⟨*proof*⟩

lemma *exp-sub-then-left-sub*:
$exp[x - (x - y)] \geq exp[y]$
⟨*proof*⟩

thm-oracles *exp-sub-then-left-sub*

Optimisations

optimization *SubAfterAddRight*: $((x + y) - y) \longmapsto x$
⟨*proof*⟩

optimization *SubAfterAddLeft*: $((x + y) - x) \longmapsto y$
⟨*proof*⟩

optimization *SubAfterSubLeft*: $((x - y) - x) \longmapsto -y$
⟨*proof*⟩

optimization *SubThenAddLeft*: $(x - (x + y)) \longmapsto -y$
⟨*proof*⟩

optimization *SubThenAddRight*: $(y - (x + y)) \longmapsto -x$
⟨*proof*⟩

optimization *SubThenSubLeft*: $(x - (x - y)) \longmapsto y$
⟨*proof*⟩

optimization *SubtractZero*: $(x - (const\ IntVal\ b\ 0)) \longmapsto x$
⟨*proof*⟩

thm-oracles *SubtractZero*

optimization *SubNegativeValue*: $(x - (-y)) \longmapsto x + y$
⟨*proof*⟩

thm-oracles *SubNegativeValue*

lemma *negate-idempotent*:
  assumes $x = IntVal\ b\ v \wedge take\text{-}bit\ b\ v = v$
  shows $x = val[-(-x)]$
  ⟨*proof*⟩

optimization *ZeroSubtractValue*: $((const\ IntVal\ b\ 0) - x) \longmapsto (-x)$
  when $(wf\text{-}stamp\ x \wedge stamp\text{-}expr\ x = IntegerStamp\ b\ lo\ hi \wedge \neg(is\text{-}ConstantExpr\ x))$
  ⟨*proof*⟩

optimization *SubSelfIsZero*: $(x - x) \longmapsto const\ IntVal\ b\ 0$ *when*
  $(wf\text{-}stamp\ x \wedge stamp\text{-}expr\ x = IntegerStamp\ b\ lo\ hi)$
  ⟨*proof*⟩

end

end

10.14 XorNode Phase

theory *XorPhase*
  imports
    Common
    Proofs.StampEvalThms
begin

phase *XorNode*
  terminating *size*
begin

lemma *bin-xor-self-is-false*:
 $bin[x \oplus x] = 0$
  ⟨*proof*⟩

lemma *bin-xor-commute*:
 $bin[x \oplus y] = bin[y \oplus x]$
  ⟨*proof*⟩

lemma *bin-eliminate-redundant-false*:
 $bin[x \oplus 0] = bin[x]$
  ⟨*proof*⟩

lemma *val-xor-self-is-false*:
  assumes $val[x \oplus x] \neq UndefVal$
  shows $val\text{-}to\text{-}bool\ (val[x \oplus x]) = False$
  ⟨*proof*⟩

lemma *val-xor-self-is-false-2*:
  assumes $(val[x \oplus x]) \neq UndefVal$
  and $x = IntVal\ 32\ v$
  shows $val[x \oplus x] = bool\text{-}to\text{-}val\ False$
  ⟨*proof*⟩

lemma *val-xor-self-is-false-3*:
  assumes $val[x \oplus x] \neq UndefVal \land x = IntVal\ 64\ v$
  shows $val[x \oplus x] = IntVal\ 64\ 0$
  ⟨*proof*⟩

lemma *val-xor-commute*:
  $val[x \oplus y] = val[y \oplus x]$
  ⟨*proof*⟩

lemma *val-eliminate-redundant-false*:
  assumes $x = new\text{-}int\ b\ v$
  assumes $val[x \oplus (bool\text{-}to\text{-}val\ False)] \neq UndefVal$
  shows $val[x \oplus (bool\text{-}to\text{-}val\ False)] = x$
  ⟨*proof*⟩

lemma *exp-xor-self-is-false*:
 assumes $wf\text{-}stamp\ x \land stamp\text{-}expr\ x = default\text{-}stamp$
 shows $exp[x \oplus x] \geq exp[false]$
  ⟨*proof*⟩

lemma *exp-eliminate-redundant-false*:
  shows $exp[x \oplus false] \geq exp[x]$
  ⟨*proof*⟩

Optimisations

optimization *XorSelfIsFalse*: $(x \oplus x) \longmapsto false\ when$
$(wf\text{-}stamp\ x \land stamp\text{-}expr\ x = default\text{-}stamp)$
  ⟨*proof*⟩

optimization *XorShiftConstantRight*: $((const\ x) \oplus y) \longmapsto y \oplus (const\ x)\ when$ $\neg(is\text{-}ConstantExpr\ y)$
  ⟨*proof*⟩

optimization *EliminateRedundantFalse*: $(x \oplus false) \longmapsto x$

⟨*proof*⟩

end

end

11 Conditional Elimination Phase

theory *ConditionalElimination*
imports
Semantics.IRTreeEvalThms
Proofs.Rewrites
Proofs.Bisimulation
begin

11.1 Individual Elimination Rules

The set of rules used for determining whether a condition $q1::'a$ implies another condition $q2::'a$ or its negation. These rules are used for conditional elimination.

inductive *impliesx* :: *IRExpr* ⇒ *IRExpr* ⇒ *bool* (- ⇛ -) **and**
impliesnot :: *IRExpr* ⇒ *IRExpr* ⇒ *bool* (- ⇛¬ -) **where**
q-imp-q:
q ⇛ *q* |
eq-impliesnot-less:
(*BinaryExpr BinIntegerEquals x y*) ⇛¬ (*BinaryExpr BinIntegerLessThan x y*) |
eq-impliesnot-less-rev:
(*BinaryExpr BinIntegerEquals x y*) ⇛¬ (*BinaryExpr BinIntegerLessThan y x*) |
less-impliesnot-rev-less:
(*BinaryExpr BinIntegerLessThan x y*) ⇛¬ (*BinaryExpr BinIntegerLessThan y x*)
|
less-impliesnot-eq:
(*BinaryExpr BinIntegerLessThan x y*) ⇛¬ (*BinaryExpr BinIntegerEquals x y*) |
less-impliesnot-eq-rev:
(*BinaryExpr BinIntegerLessThan x y*) ⇛¬ (*BinaryExpr BinIntegerEquals y x*) |
negate-true:
⟦*x* ⇛¬ *y*⟧ ⟹ *x* ⇛ (*UnaryExpr UnaryLogicNegation y*) |
negate-false:
⟦*x* ⇛ *y*⟧ ⟹ *x* ⇛¬ (*UnaryExpr UnaryLogicNegation y*)

The relation *q1*::*IRExpr* ⇛ *q2*::*IRExpr* indicates that the implication (*q1*::*bool*) ⟶ (*q2*::*bool*) is known true (i.e. universally valid), and the relation *q1*::*IRExpr* ⇛¬ *q2*::*IRExpr* indicates that the implication (*q1*::*bool*) ⟶ (*q2*::*bool*) is known false (i.e. (*q1*::*bool*) ⟶ ¬ (*q2*::*bool*) is universally valid. If neither

$q1::IRExpr \Rrightarrow q2::IRExpr$ nor $q1::IRExpr \Rrightarrow\neg\ q2::IRExpr$ then the status is unknown. Only the known true and known false cases can be used for conditional elimination.

fun *implies-valid* :: *IRExpr* ⇒ *IRExpr* ⇒ *bool* (**infix** ↣ *50*) **where**
  implies-valid q1 q2 =
    (∀ *m p v1 v2*. ([*m*, *p*] ⊢ *q1* ↦ *v1*) ∧ ([*m*,*p*] ⊢ *q2* ↦ *v2*) ⟶
        (*val-to-bool v1* ⟶ *val-to-bool v2*))

fun *impliesnot-valid* :: *IRExpr* ⇒ *IRExpr* ⇒ *bool* (**infix** ⤔ *50*) **where**
  impliesnot-valid q1 q2 =
    (∀ *m p v1 v2*. ([*m*, *p*] ⊢ *q1* ↦ *v1*) ∧ ([*m*,*p*] ⊢ *q2* ↦ *v2*) ⟶
        (*val-to-bool v1* ⟶ ¬*val-to-bool v2*))

The relation $(q1::IRExpr) \rightarrowtail (q2::IRExpr)$ means $(q1::bool) \longrightarrow (q2::bool)$ is universally valid, and the relation $(q1::IRExpr)$ ⤔ $(q2::IRExpr)$ means $(q1::bool) \longrightarrow \neg\ (q2::bool)$ is universally valid.

lemma *eq-impliesnot-less-helper*:
  v1 = *v2* ⟶ ¬(*int-signed-value b v1* < *int-signed-value b v2*)
  ⟨*proof*⟩

lemma *eq-impliesnot-less-val*:
  val-to-bool(*intval-equals v1 v2*) ⟶ ¬*val-to-bool*(*intval-less-than v1 v2*)
  ⟨*proof*⟩

lemma *eq-impliesnot-less-rev-val*:
  val-to-bool(*intval-equals v1 v2*) ⟶ ¬*val-to-bool*(*intval-less-than v2 v1*)
⟨*proof*⟩

lemma *less-impliesnot-rev-less-val*:
  val-to-bool(*intval-less-than v1 v2*) ⟶ ¬*val-to-bool*(*intval-less-than v2 v1*)
  ⟨*proof*⟩

lemma *less-impliesnot-eq-val*:
  val-to-bool(*intval-less-than v1 v2*) ⟶ ¬*val-to-bool*(*intval-equals v1 v2*)
  ⟨*proof*⟩

lemma *logic-negate-type*:
  assumes [*m*, *p*] ⊢ *UnaryExpr UnaryLogicNegation x* ↦ *v*
  shows ∃ *b v2*. [*m*, *p*] ⊢ *x* ↦ *IntVal b v2*
  ⟨*proof*⟩

lemma *intval-logic-negation-inverse*:
  assumes *b* > *0*
  assumes *x* = *IntVal b v*
  shows *val-to-bool* (*intval-logic-negation x*) ⟷ ¬(*val-to-bool x*)
  ⟨*proof*⟩

lemma *logic-negation-relation-tree*:

assumes $[m,\ p] \vdash y \mapsto \mathit{val}$
assumes $[m,\ p] \vdash \mathit{UnaryExpr\ UnaryLogicNegation\ y} \mapsto \mathit{invval}$
shows $\mathit{val\text{-}to\text{-}bool\ val} \longleftrightarrow \neg(\mathit{val\text{-}to\text{-}bool\ invval})$
$\langle \mathit{proof} \rangle$

The following theorem shows that the known true/false rules are valid.

theorem *implies-impliesnot-valid*:
shows $((q1 \Rrightarrow q2) \longrightarrow (q1 \rightarrowtail q2)) \wedge$
$((q1 \Rrightarrow\neg\ q2) \longrightarrow (q1 \nrightarrowtail q2))$
(**is** $(?imp \longrightarrow ?val) \wedge (?notimp \longrightarrow ?notval)$)
$\langle \mathit{proof} \rangle$

We introduce a type $\mathit{TriState}::'a$ (as in the GraalVM compiler) to represent when static analysis can tell us information about the value of a Boolean expression. If $\mathit{Unknown}::'a$ then no information can be inferred and if $\mathit{KnownTrue}::'a/\mathit{KnownFalse}::'a$ one can infer the expression is always true/false.

datatype $\mathit{TriState} = \mathit{Unknown} \mid \mathit{KnownTrue} \mid \mathit{KnownFalse}$

The implies relation corresponds to the LogicNode.implies method from the compiler which attempts to infer when one logic nodes value can be inferred from a known logic node.

inductive *implies* :: $\mathit{IRGraph} \Rightarrow \mathit{IRNode} \Rightarrow \mathit{IRNode} \Rightarrow \mathit{TriState} \Rightarrow \mathit{bool}$
(- $\vdash$ - & - $\hookrightarrow$ -) **for** g **where**
eq-imp-less:
$g \vdash (\mathit{IntegerEqualsNode}\ x\ y)\ \&\ (\mathit{IntegerLessThanNode}\ x\ y) \hookrightarrow \mathit{KnownFalse} \mid$
eq-imp-less-rev:
$g \vdash (\mathit{IntegerEqualsNode}\ x\ y)\ \&\ (\mathit{IntegerLessThanNode}\ y\ x) \hookrightarrow \mathit{KnownFalse} \mid$
less-imp-rev-less:
$g \vdash (\mathit{IntegerLessThanNode}\ x\ y)\ \&\ (\mathit{IntegerLessThanNode}\ y\ x) \hookrightarrow \mathit{KnownFalse} \mid$
less-imp-not-eq:
$g \vdash (\mathit{IntegerLessThanNode}\ x\ y)\ \&\ (\mathit{IntegerEqualsNode}\ x\ y) \hookrightarrow \mathit{KnownFalse} \mid$
less-imp-not-eq-rev:
$g \vdash (\mathit{IntegerLessThanNode}\ x\ y)\ \&\ (\mathit{IntegerEqualsNode}\ y\ x) \hookrightarrow \mathit{KnownFalse} \mid$

x-imp-x:
$g \vdash x\ \&\ x \hookrightarrow \mathit{KnownTrue} \mid$

negate-false:
$[\![g \vdash x\ \&\ (\mathit{kind}\ g\ y) \hookrightarrow \mathit{KnownTrue}]\!] \Longrightarrow g \vdash x\ \&\ (\mathit{LogicNegationNode}\ y) \hookrightarrow \mathit{KnownFalse} \mid$
negate-true:
$[\![g \vdash x\ \&\ (\mathit{kind}\ g\ y) \hookrightarrow \mathit{KnownFalse}]\!] \Longrightarrow g \vdash x\ \&\ (\mathit{LogicNegationNode}\ y) \hookrightarrow \mathit{KnownTrue}$

Total relation over partial implies relation

inductive *condition-implies* :: $\mathit{IRGraph} \Rightarrow \mathit{IRNode} \Rightarrow \mathit{IRNode} \Rightarrow \mathit{TriState} \Rightarrow \mathit{bool}$
(- $\vdash$ - & - $\rightharpoonup$ -) **for** g **where**
$[\![\neg(g \vdash a\ \&\ b \hookrightarrow \mathit{imp})]\!] \Longrightarrow (g \vdash a\ \&\ b \rightharpoonup \mathit{Unknown}) \mid$

$[\![(g \vdash a \ \& \ b \hookrightarrow imp)]\!] \Longrightarrow (g \vdash a \ \& \ b \rightharpoonup imp)$

inductive *implies-tree* :: *IRExpr* ⇒ *IRExpr* ⇒ *bool* ⇒ *bool*
  (- & - ↪ -) **where**
  eq-imp-less:
  (*BinaryExpr BinIntegerEquals x y*) & (*BinaryExpr BinIntegerLessThan x y*) ↪ *False* |
  eq-imp-less-rev:
  (*BinaryExpr BinIntegerEquals x y*) & (*BinaryExpr BinIntegerLessThan y x*) ↪ *False* |
  less-imp-rev-less:
  (*BinaryExpr BinIntegerLessThan x y*) & (*BinaryExpr BinIntegerLessThan y x*) ↪ *False* |
  less-imp-not-eq:
  (*BinaryExpr BinIntegerLessThan x y*) & (*BinaryExpr BinIntegerEquals x y*) ↪ *False* |
  less-imp-not-eq-rev:
  (*BinaryExpr BinIntegerLessThan x y*) & (*BinaryExpr BinIntegerEquals y x*) ↪ *False* |
  x-imp-x:
  x & *x* ↪ *True* |
  negate-false:
  ⟦*x* & *y* ↪ *True*⟧ ⟹ *x* & (*UnaryExpr UnaryLogicNegation y*) ↪ *False* |
  negate-true:
  ⟦*x* & *y* ↪ *False*⟧ ⟹ *x* & (*UnaryExpr UnaryLogicNegation y*) ↪ *True*

Proofs that the implies relation is correct with respect to the existing evaluation semantics.

lemma *logic-negation-relation*:
  assumes [*g*, *m*, *p*] ⊢ *y* ↦ *val*
  assumes *kind g neg* = *LogicNegationNode y*
  assumes [*g*, *m*, *p*] ⊢ *neg* ↦ *invval*
  assumes *invval* ≠ *UndefVal*
  shows *val-to-bool val* ⟷ ¬(*val-to-bool invval*)
  ⟨*proof*⟩

lemma *implies-valid*:
  assumes *x* & *y* ↪ *imp*
  assumes [*m*, *p*] ⊢ *x* ↦ *v1*
  assumes [*m*, *p*] ⊢ *y* ↦ *v2*
  shows (*imp* ⟶ (*val-to-bool v1* ⟶ *val-to-bool v2*)) ∧
        (¬*imp* ⟶ (*val-to-bool v1* ⟶ ¬(*val-to-bool v2*)))
    (**is** (*?TP* ⟶ *?TC*) ∧ (*?FP* ⟶ *?FC*))
  ⟨*proof*⟩

lemma *implies-true-valid*:
  assumes *x* & *y* ↪ *imp*

assumes imp
assumes $[m, p] \vdash x \mapsto v1$
assumes $[m, p] \vdash y \mapsto v2$
shows $val\text{-}to\text{-}bool\ v1 \longrightarrow val\text{-}to\text{-}bool\ v2$
⟨*proof*⟩

lemma *implies-false-valid*:
assumes $x\ \&\ y \hookrightarrow imp$
assumes $\neg imp$
assumes $[m, p] \vdash x \mapsto v1$
assumes $[m, p] \vdash y \mapsto v2$
shows $val\text{-}to\text{-}bool\ v1 \longrightarrow \neg(val\text{-}to\text{-}bool\ v2)$
⟨*proof*⟩

The following relation corresponds to the UnaryOpLogicNode.tryFold and BinaryOpLogicNode.tryFold methods and their associated concrete implementations.

The relation determines if a logic operation can be shown true or false through the stamp typing information.

inductive $tryFold :: IRNode \Rightarrow (ID \Rightarrow Stamp) \Rightarrow bool \Rightarrow bool$
where
$[\![alwaysDistinct\ (stamps\ x)\ (stamps\ y)]\!]$
$\Longrightarrow tryFold\ (IntegerEqualsNode\ x\ y)\ stamps\ False\ |$
$[\![neverDistinct\ (stamps\ x)\ (stamps\ y)]\!]$
$\Longrightarrow tryFold\ (IntegerEqualsNode\ x\ y)\ stamps\ True\ |$
$[\![is\text{-}IntegerStamp\ (stamps\ x);$
$is\text{-}IntegerStamp\ (stamps\ y);$
$stpi\text{-}upper\ (stamps\ x) < stpi\text{-}lower\ (stamps\ y)]\!]$
$\Longrightarrow tryFold\ (IntegerLessThanNode\ x\ y)\ stamps\ True\ |$
$[\![is\text{-}IntegerStamp\ (stamps\ x);$
$is\text{-}IntegerStamp\ (stamps\ y);$
$stpi\text{-}lower\ (stamps\ x) \geq stpi\text{-}upper\ (stamps\ y)]\!]$
$\Longrightarrow tryFold\ (IntegerLessThanNode\ x\ y)\ stamps\ False$

Proofs that show that when the stamp lookup function is well-formed, the tryFold relation correctly predicts the output value with respect to our evaluation semantics.

lemma
assumes $kind\ g\ nid = IntegerEqualsNode\ x\ y$
assumes $[g, m, p] \vdash nid \mapsto v$
assumes $([g, m, p] \vdash x \mapsto xval) \wedge ([g, m, p] \vdash y \mapsto yval)$
shows $val\text{-}to\text{-}bool\ (intval\text{-}equals\ xval\ yval) \longleftrightarrow v = IntVal\ 32\ 1$
⟨*proof*⟩

lemma *tryFoldIntegerEqualsAlwaysDistinct*:
assumes $wf\text{-}stamp\ g\ stamps$
assumes $kind\ g\ nid = (IntegerEqualsNode\ x\ y)$
assumes $[g, m, p] \vdash nid \mapsto v$

assumes *alwaysDistinct* (*stamps x*) (*stamps y*)
shows $v = \mathit{IntVal}\ 32\ 0$
⟨*proof*⟩

lemma *tryFoldIntegerEqualsNeverDistinct*:
assumes *wf-stamp g stamps*
assumes *kind g nid* = (*IntegerEqualsNode x y*)
assumes $[g,\ m,\ p] \vdash nid \mapsto v$
assumes *neverDistinct* (*stamps x*) (*stamps y*)
shows $v = \mathit{IntVal}\ 32\ 1$
⟨*proof*⟩

lemma *tryFoldIntegerLessThanTrue*:
assumes *wf-stamp g stamps*
assumes *kind g nid* = (*IntegerLessThanNode x y*)
assumes $[g,\ m,\ p] \vdash nid \mapsto v$
assumes *stpi-upper* (*stamps x*) < *stpi-lower* (*stamps y*)
shows $v = \mathit{IntVal}\ 32\ 1$
⟨*proof*⟩

lemma *tryFoldIntegerLessThanFalse*:
assumes *wf-stamp g stamps*
assumes *kind g nid* = (*IntegerLessThanNode x y*)
assumes $[g,\ m,\ p] \vdash nid \mapsto v$
assumes *stpi-lower* (*stamps x*) ≥ *stpi-upper* (*stamps y*)
shows $v = \mathit{IntVal}\ 32\ 0$
⟨*proof*⟩

theorem *tryFoldProofTrue*:
assumes *wf-stamp g stamps*
assumes *tryFold* (*kind g nid*) *stamps True*
assumes $[g,\ m,\ p] \vdash nid \mapsto v$
shows *val-to-bool v*
⟨*proof*⟩

theorem *tryFoldProofFalse*:
assumes *wf-stamp g stamps*
assumes *tryFold* (*kind g nid*) *stamps False*
assumes $[g,\ m,\ p] \vdash nid \mapsto v$
shows ¬(*val-to-bool v*)
⟨*proof*⟩

inductive-cases *StepE*:
$g,\ p \vdash (nid,m,h) \rightarrow (nid',m',h)$

Perform conditional elimination rewrites on the graph for a particular node. In order to determine conditional eliminations appropriately the rule needs two data structures produced by static analysis. The first parameter is the set of IRNodes that we know result in a true value when evaluated. The second parameter is a mapping from node identifiers to the flow-sensitive stamp.

The relation transforms the third parameter to the fifth parameter for a node identifier which represents the fourth parameter.

inductive *ConditionalEliminationStep* ::
$IRExpr\ set \Rightarrow (ID \Rightarrow Stamp) \Rightarrow IRGraph \Rightarrow ID \Rightarrow IRGraph \Rightarrow bool$ **where**
impliesTrue:
$[\![kind\ g\ ifcond = (IfNode\ cid\ t\ f);$
$g \vdash cid \simeq cond;$
$\exists\ ce \in conds\ .\ (ce \Rrightarrow cond);$
$g' = constantCondition\ True\ ifcond\ (kind\ g\ ifcond)\ g$
$]\!] \Longrightarrow ConditionalEliminationStep\ conds\ stamps\ g\ ifcond\ g'$ |

impliesFalse:
$[\![kind\ g\ ifcond = (IfNode\ cid\ t\ f);$
$g \vdash cid \simeq cond;$
$\exists\ ce \in conds\ .\ (ce \Rrightarrow\neg\ cond);$
$g' = constantCondition\ False\ ifcond\ (kind\ g\ ifcond)\ g$
$]\!] \Longrightarrow ConditionalEliminationStep\ conds\ stamps\ g\ ifcond\ g'$ |

tryFoldTrue:
$[\![kind\ g\ ifcond = (IfNode\ cid\ t\ f);$
$cond = kind\ g\ cid;$
$tryFold\ (kind\ g\ cid)\ stamps\ True;$
$g' = constantCondition\ True\ ifcond\ (kind\ g\ ifcond)\ g$
$]\!] \Longrightarrow ConditionalEliminationStep\ conds\ stamps\ g\ ifcond\ g'$ |

tryFoldFalse:
$[\![kind\ g\ ifcond = (IfNode\ cid\ t\ f);$
$cond = kind\ g\ cid;$
$tryFold\ (kind\ g\ cid)\ stamps\ False;$
$g' = constantCondition\ False\ ifcond\ (kind\ g\ ifcond)\ g$
$]\!] \Longrightarrow ConditionalEliminationStep\ conds\ stamps\ g\ ifcond\ g'$

code-pred (*modes*: $i \Rightarrow i \Rightarrow i \Rightarrow i \Rightarrow o \Rightarrow bool$) *ConditionalEliminationStep*
⟨*proof*⟩

thm *ConditionalEliminationStep.equation*

11.2 Control-flow Graph Traversal

type-synonym *Seen* = *ID set*
type-synonym *Condition* = *IRExpr*
type-synonym *Conditions* = *Condition list*
type-synonym *StampFlow* = $(ID \Rightarrow Stamp)$ *list*

nextEdge helps determine which node to traverse next by returning the first successor edge that isn't in the set of already visited nodes. If there is not an appropriate successor, None is returned instead.

fun *nextEdge* :: *Seen* ⇒ *ID* ⇒ *IRGraph* ⇒ *ID option* **where**
  nextEdge seen nid g =
    (*let nids* = (*filter* (λ*nid'*. *nid'* ∉ *seen*) (*successors-of* (*kind g nid*))) *in*
     (*if length nids* > *0 then Some* (*hd nids*) *else None*))

pred determines which node, if any, acts as the predecessor of another.

Merge nodes represent a special case where-in the predecessor exists as an input edge of the merge node, to simplify the traversal we treat only the first input end node as the predecessor, ignoring that multiple nodes may act as a successor.

For all other nodes, the predecessor is the first element of the predecessors set. Note that in a well-formed graph there should only be one element in the predecessor set.

fun *pred* :: *IRGraph* ⇒ *ID* ⇒ *ID option* **where**
  pred g nid = (*case kind g nid of*
    (*MergeNode ends - -*) ⇒ *Some* (*hd ends*) |
    - ⇒
      (*if IRGraph.predecessors g nid* = {}
        then None else
        Some (*hd* (*sorted-list-of-set* (*IRGraph.predecessors g nid*)))
)
)

When the basic block of an if statement is entered, we know that the condition of the preceding if statement must be true. As in the GraalVM compiler, we introduce the registerNewCondition funciton which roughly corresponds to the ConditionalEliminationPhase.registerNewCondition. This method updates the flow-sensitive stamp information based on the condition which we know must be true.

fun *clip-upper* :: *Stamp* ⇒ *int* ⇒ *Stamp* **where**
  clip-upper (*IntegerStamp b l h*) *c* = (*IntegerStamp b l c*) |
  clip-upper s c = *s*
fun *clip-lower* :: *Stamp* ⇒ *int* ⇒ *Stamp* **where**
  clip-lower (*IntegerStamp b l h*) *c* = (*IntegerStamp b c h*) |
  clip-lower s c = *s*

fun *registerNewCondition* :: *IRGraph* ⇒ *IRNode* ⇒ (*ID* ⇒ *Stamp*) ⇒ (*ID* ⇒ *Stamp*) **where**

  registerNewCondition g (*IntegerEqualsNode x y*) *stamps* =
    (*stamps*
      (*x* := *join* (*stamps x*) (*stamps y*)))
      (*y* := *join* (*stamps x*) (*stamps y*)) |

registerNewCondition g (*IntegerLessThanNode x y*) *stamps* =
(*stamps*
(*x* := *clip-upper* (*stamps x*) (*stpi-lower* (*stamps y*))))
(*y* := *clip-lower* (*stamps y*) (*stpi-upper* (*stamps x*))) |
registerNewCondition g - *stamps* = *stamps*

fun *hdOr* :: $'a$ *list* $\Rightarrow$ $'a \Rightarrow 'a$ **where**
hdOr (*x* # *xs*) *de* = *x* |
hdOr [] *de* = *de*

The Step relation is a small-step traversal of the graph which handles transitions between individual nodes of the graph.

It relates a pairs of tuple of the current node, the set of seen nodes, the always true stack of IfNode conditions, and the flow-sensitive stamp information.

inductive *Step*
:: *IRGraph* $\Rightarrow$ (*ID* $\times$ *Seen* $\times$ *Conditions* $\times$ *StampFlow*) $\Rightarrow$ (*ID* $\times$ *Seen* $\times$ *Conditions* $\times$ *StampFlow*) *option* $\Rightarrow$ *bool*
for *g* **where**
— Hit a BeginNode with an IfNode predecessor which represents the start of a basic block for the IfNode. 1. nid' will be the successor of the begin node. 2. Find the first and only predecessor. 3. Extract condition from the preceding IfNode. 4. Negate condition if the begin node is second branch (we've taken the else branch of the condition) 5. Add the condition or the negated condition to stack 6. Perform any stamp updates based on the condition using the registerNewCondition function and place them on the top of the stack of stamp information
⟦*kind g nid* = *BeginNode nid'*;

nid $\notin$ *seen*;
seen' = {*nid*} $\cup$ *seen*;

Some ifcond = *pred g nid*;
kind g ifcond = *IfNode cond t f*;

i = *find-index nid* (*successors-of* (*kind g ifcond*));
c = (*if i* = *0 then kind g cond else LogicNegationNode cond*);
rep g cond ce;
ce' = (*if i* = *0 then ce else UnaryExpr UnaryLogicNegation ce*);
conds' = *ce'* # *conds*;

flow' = *registerNewCondition g c* (*hdOr flow* (*stamp g*))⟧
$\Longrightarrow$ *Step g* (*nid*, *seen*, *conds*, *flow*) (*Some* (*nid'*, *seen'*, *conds'*, *flow'* # *flow*)) |

— Hit an EndNode 1. nid' will be the usage of EndNode 2. pop the conditions and stamp stack
⟦*kind g nid* = *EndNode*;

nid $\notin$ *seen*;
seen' = {*nid*} $\cup$ *seen*;

$nid' = any\text{-}usage\ g\ nid;$

$conds' = tl\ conds;$
$flow' = tl\ flow\rrbracket$
$\Longrightarrow Step\ g\ (nid, seen, conds, flow)\ (Some\ (nid', seen', conds', flow'))\ |$

— We can find a successor edge that is not in seen, go there
$\llbracket\neg(is\text{-}EndNode\ (kind\ g\ nid));$
$\neg(is\text{-}BeginNode\ (kind\ g\ nid));$

$nid \notin seen;$
$seen' = \{nid\} \cup seen;$

$Some\ nid' = nextEdge\ seen'\ nid\ g\rrbracket$
$\Longrightarrow Step\ g\ (nid, seen, conds, flow)\ (Some\ (nid', seen', conds, flow))\ |$

— We can cannot find a successor edge that is not in seen, give back None
$\llbracket\neg(is\text{-}EndNode\ (kind\ g\ nid));$
$\neg(is\text{-}BeginNode\ (kind\ g\ nid));$

$nid \notin seen;$
$seen' = \{nid\} \cup seen;$

$None = nextEdge\ seen'\ nid\ g\rrbracket$
$\Longrightarrow Step\ g\ (nid, seen, conds, flow)\ None\ |$

— We've already seen this node, give back None
$\llbracket nid \in seen\rrbracket \Longrightarrow Step\ g\ (nid, seen, conds, flow)\ None$

code-pred $(modes: i \Rightarrow i \Rightarrow o \Rightarrow bool)\ Step\ \langle proof\rangle$

The ConditionalEliminationPhase relation is responsible for combining the individual traversal steps from the Step relation and the optimizations from the ConditionalEliminationStep relation to perform a transformation of the whole graph.

inductive $ConditionalEliminationPhase$
$:: IRGraph \Rightarrow (ID \times Seen \times Conditions \times StampFlow) \Rightarrow IRGraph \Rightarrow bool$
where

— Can do a step and optimise for the current node
$\llbracket Step\ g\ (nid, seen, conds, flow)\ (Some\ (nid', seen', conds', flow'));$
$ConditionalEliminationStep\ (set\ conds)\ (hdOr\ flow\ (stamp\ g))\ g\ nid\ g';$

$ConditionalEliminationPhase\ g'\ (nid', seen', conds', flow')\ g''\rrbracket$
$\Longrightarrow ConditionalEliminationPhase\ g\ (nid, seen, conds, flow)\ g''\ |$

— Can do a step, matches whether optimised or not causing non-determinism We need to find a way to negate ConditionalEliminationStep

$[\![Step\ g\ (nid,\ seen,\ conds,\ flow)\ (Some\ (nid',\ seen',\ conds',\ flow'));$

$ConditionalEliminationPhase\ g\ (nid',\ seen',\ conds',\ flow')\ g']\!]$
$\Longrightarrow ConditionalEliminationPhase\ g\ (nid,\ seen,\ conds,\ flow)\ g'\ |$

— Can't do a step but there is a predecessor we can backtrace to
$[\![Step\ g\ (nid,\ seen,\ conds,\ flow)\ None;$
$Some\ nid' = pred\ g\ nid;$
$seen' = \{nid\} \cup seen;$
$ConditionalEliminationPhase\ g\ (nid',\ seen',\ conds,\ flow)\ g']\!]$
$\Longrightarrow ConditionalEliminationPhase\ g\ (nid,\ seen,\ conds,\ flow)\ g'\ |$

— Can't do a step and have no predecessors so terminate
$[\![Step\ g\ (nid,\ seen,\ conds,\ flow)\ None;$
$None = pred\ g\ nid]\!]$
$\Longrightarrow ConditionalEliminationPhase\ g\ (nid,\ seen,\ conds,\ flow)\ g$

code-pred (*modes*: $i \Rightarrow i \Rightarrow o \Rightarrow bool$) *ConditionalEliminationPhase* ⟨*proof*⟩

definition *runConditionalElimination* :: *IRGraph* ⇒ *IRGraph* **where**
runConditionalElimination g =
(*Predicate.the* (*ConditionalEliminationPhase-i-i-o g* (*0*, {}, ([], []))))

end