# Software Engineering using Formal Methods Proof Obligations

Wolfgang Ahrendt, Mauricio Chimento

14 October 2016

## This Part

making the connection between

**JML** 

and

Dynamic Logic / KeY

- generating,
- understanding,
- and proving

DL proof obligations from JML specifications

# From JML Contracts to Intermediate Format to Proof Obligations (PO)

```
public class A {
   /*@ public normal_behavior
   @ requires <Precondition>;
   @ ensures <Postcondition>;
   @ assignable <locations>;
   @*/
public int m(params) {..}
}
Intermediate Format

(pre, post, div, var, mod)

Translation

PO Generation

PO Generation

PO Generation
```

Proof obligation as DL formula

```
pre \rightarrow \\ \langle \texttt{this.m(params);} \rangle \\ (post \& frame)
```

## JML Translation: Normalizing JML Contracts

#### Normalization of JML Contracts

- 1. Flattening of nested specifications
- 2. Making implicit specifications explicit
- 3. Processing of modifiers
- 4. Adding of default clauses if not present
- 5. Contraction of several clauses

Tho following introduces principles of this process

nested specification cases allow to factor out common preconditions

```
@ public normal_behavior
@ requires R;
0 { |
@
    requires R1;
0
    ensures E1;
@
    assignable A1;
@
0
    also
0
    requires R2;
0
    ensures E2;
0
    assignable A2;
  1}
expands to ... (next page)
```

```
(previous page) ... expands to
@ public normal_behavior
0 requires R;
@ requires R1;
@ ensures E1:
@ assignable A1;
0
 also
0
@ public normal_behavior
@ requires R;
@ requires R2;
@ ensures E2;
@ assignable A2;
```

```
@ public normal_behavior
@ requires amount > 0;
@ { |
0
    requires amount + balance < limit && isValid()==true;</pre>
0
    ensures \result == true;
0
    ensures balance == amount + \old(balance);
@
    assignable balance;
0
0
    also
@
0
    requires amount + balance >= limit;
0
    ensures \result == false;
0
    ensures unsuccessfulOperations
             == \old(unsuccessfulOperations) + 1;
@
@
    assignable unsuccessfulOperations;
expands to ... (next page)
```

```
(previous page) ... expands to
@ public normal_behavior
@ requires amount > 0;
@ requires amount + balance < limit && isValid() == true;
@ ensures \result == true;
@ ensures balance == amount + \old(balance);
@ assignable balance;
0
@ also
0
@ public normal behavior
@ requires amount > 0;
@ requires amount + balance >= limit;
@ ensures \result == false;
@ ensures unsuccessfulOperations
0
          == \old(unsuccessfulOperations) + 1;
  assignable unsuccessfulOperations;
```

SEFM: Proof Obligations

## **Normalisation:**

# **Making Implicit Specifications Explicit**

## **Implicit Specifications**

- Kind of behavior
- non\_null by default
- Implicit \invariant\_for(this) in requires, ensures & signals clause

#### Making 'kind of behavior' explicit

- Deactivate implicit behavior specification: replace normal\_behavior/exceptional\_behavior by behavior
- 2. Add in case of replaced
  - normal\_behavior the clause signals (Throwable t) false;
  - exceptional\_behavior the clause ensures false;

### **Normalisation:**

## **Making Implicit Specifications Explicit**

## **Implicit Specifications**

- Kind of behavior
- ▶ non\_null by default
- Implicit \invariant\_for(this) as requires, ensures & signals
  clause

## Making non\_null explicit for method specifications

- 1. Where nullable is absent, forbid null through preconditions (for parameters<sup>a</sup>) and postcondition (for return value<sup>a</sup>).
  - E.g., for method void m(Object o) add requires o != null;
- 2. Deactivate implicit non\_null by adding nullable, where absent, to parameters<sup>a</sup> and return type declarations<sup>a</sup>

<sup>&</sup>lt;sup>a</sup>reference typed

## **Normalisation:**

# **Making Implicit Specifications Explicit**

#### **Implicit Specifications**

- Kind of behavior
- ▶ non\_null by default
- Implicit \invariant\_for(this) as requires, ensures & signals
  clause

#### Making \invariant\_for(this) explicit for method specifications

- 1. Add explicit \invariant\_for(this) to non-helper method specs, as
  - requires \invariant\_for(this);
  - ensures \invariant\_for(this);
  - ▶ signals (Throwable t) \invariant\_for(this);
- Deactivate implicit \invariant\_for(this) by adding helper modifier to method (if not already present)

```
/*@ public normal_behavior
  @ requires c.id >= 0;
  @ ensures \result == ( ... );
  @*/
  public boolean addCategory(Category c) {
```

Kind of behavior

```
/*@ public behavior
  @ requires c.id >= 0;
  @ ensures \result == ( ... );
  @ signals (Throwable exc) false;
  @*/
  public boolean addCategory(Category c) {
```

```
non_null by default

/*@ public behavior
@ requires c.id >= 0;
@ requires c != null;
@ ensures \result == ( ... );
@ signals (Throwable exc) false;}
@*/
public boolean addCategory(/*@ nullable @*/ Category c) {
```

Implicit \invariant\_for(this) as requires, ensures & signals
clause

```
/*@ public behavior
@ requires c.id >= 0;
@ requires c != null;
@ requires \invariant_for(this);
@ ensures \result == ( ... );
@ ensures \invariant_for(this);
@ signals (Throwable exc) false;
@ signals (Throwable exc) \invariant_for(this);
@*/
public /*@ helper @*/
boolean addCategory(/*@ nullable @*/ Category c) {
```

Implicit specification explicit

```
/*@ public behavior
@ requires c.id >= 0;
@ requires c != null;
@ requires \invariant_for(this);
@ ensures \result == ( ... );
@ ensures \invariant_for(this);
@ signals (Throwable exc) false;
@ signals (Throwable exc) \invariant_for(this);
@*/
public /*@ helper @*/
boolean addCategory(/*@ nullable @*/ Category c) {
```

## Normalisation: Processing of Modifiers

## Processing of Modifiers (Not detailed)

Expanding pure modifier: add to each specification case:

- ► assignable \nothing;
- ▶ diverges false;

#### Adding of default clauses if not present

Where clauses with defaults (e.g., diverges, assignable) are absent, add explicit clauses.

### **Normalisation: Clause Contraction**

#### Clause Contraction

Merge multiple clauses of the same kind into a single one of that kind.

For instance,

```
/*@ public behavior
                            /*@ public behavior
 @ requires R1;
                              @ requires R1 && R2;
 @ requires R2;
                              @ ensures E1 && E2;
 @ ensures E1;
                              @ signals (Throwable exc)
                                 (exc instanceof T1 ==> S1)
 @ ensures E2;
 @ signals (T1 exc) S1;
                                 &r.&r.
 @ signals (T2 exc) S2:
                              @ (exc instanceof T2 ==> S2):
 @*/
                              0*/
```

#### Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- a precondition DL formula pre,
- a postcondition DL formula post,
- → a divergence indicator div ∈ {TOTAL, PARTIAL},
- a variant var a term of type any
- ▶ a modifies set mod, either of type LocSet or \strictly\_nothing

SEFM: Proof Obligations CHALMERS/GU 161014 18 / 49

# Translating JML Expressions to DL-Terms: Arithmetic Expressions

Translation replaces arithmetic JAVA operators by generalized operators Generic towards various integer semantics (JAVA, Math).

#### Example:

```
"+" becomes "javaAddInt" or "javaAddLong"
"-" becomes "javaSubInt" or "javaSubLong"
...
```

SEFM: Proof Obligations

# Translating JML Expressions to DL-Terms: The this Reference

The this reference, explicit or implicit, has only a meaning within a program (refers to currently executing instance).

On logic level (outside the modalities) no such context exists.

```
this reference translated to a program variable (named by convention)
self
```

```
e.g., given class
public class MyClass {
   int f;
}
JML expressions f and this.f
translated to
DL term select(heap, self, f)
```

# **Translating Boolean JML Expressions**

First-order logic treated fundamentally different in JML and KeY logic

#### **JML**

- Formulas no separate syntactic category
- Instead: JAVA's boolean expressions extended with first-order concepts (i.p. quantifiers)

#### **Dynamic Logic**

- ► Formulas and expressions completely separate
- ► Truth constants true, false are formulas, boolean constants TRUE, FALSE are terms
- ► Atomic formulas take terms as arguments; e.g.:
  - x y < 5</p>
  - b = TRUE

## **Translating Boolean JML Expressions**

v/f/m() boolean variables/fields/pure methods b\_0, b\_1 boolean JML expressions, e\_0, e\_1 JML expressions  $\mathcal E$  translates JML expressions to DL terms

# ${\mathcal F}$ Translates boolean JML Expressions to Formulas

Quantified formulas over reference types:

```
  \mathcal{F}((\forall\ T\ x;\ e\_0;\ e\_1)) = \\  \forall\ T\ x;\ ( \\  \  \  (!x=null\ \&\ select(heap,x,\created>)=TRUE\ \&\ \mathcal{F}(e\_0)) \\ -> \mathcal{F}(e\_1))    \mathcal{F}((\ext{lexists}\ T\ x;\ e\_0;\ e\_1)) = \\  \forall\ \&\ select(heap,x,\created>)=TRUE\ \&\ \mathcal{F}(e\_0)) \\  \ \&\ \mathcal{F}(e\_1))
```

# $\mathcal F$ Translates boolean JML Expressions to Formulas

Quantified formulas over primitive types, e.g., int

inInt (similar inLong, inByte):

Predefined predicate symbol with fixed interpretation

Meaning: Argument is within the range of the Java int datatype.

# **Translating Class Invariants**

```
\mathcal{F}(\text{invariant\_for(e)}) = \text{Object} ::< inv>(heap, \mathcal{E}(e))
```

- \invariant\_for(e) translated to built-in predicate Object ::<inv>,
   applied to heap and the translation of e
- Object ::<inv> is considered a specification-only field <inv> of class Object (inherited by all sub-types of Object)
- ▶ Given that o is of type T, KeY can expand Object ::<inv>(heap, o) to the invariant of T
- ► Object ::<inv>(heap, o) pretty printed as o.<inv>()
- ▶ Read 'invariant of o'

#### Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- a precondition DL formula pre
- ▶ a postcondition DL formula post ✓?almost,
- ▶ a divergence indicator  $div \in \{TOTAL, PARTIAL\}$ ,
- a variant var a term of type any,
- ▶ a modifies set mod, either of type LocSet or \strictly\_nothing

SEFM: Proof Obligations

### **Translation of Ensures Clauses**

What is missing for ensures clauses?

- ► Translation of \result
- ► Translation of \old(.) expressions

SEFM: Proof Obligations CHALMERS/GU 161014 26 / 49

## **Translation of Ensures Clauses**

#### Translating \result

For \result used in ensures clause of method T m(...):

$$\mathcal{E}(\text{result}) = \text{result}$$

where result  $\in PVar$  of type T does not occur in the program.

# Translating \old Expressions

 $\old(e)$  evaluates e in the prestate of the method Accesses to heap must be evaluated w.r.t. the 'old' heap

- Introduce a global program variables heapAtPre of type Heap (Intention: heapAtPre refers to heap in method's pre-state)
- 2. Define:

$$\mathcal{E}(\text{\ensuremath{$\setminus$}}(e)) = \mathcal{E}_{ ensuremath{$\setminus$}}^{ ensuremath{$\setminus$}}(e)$$
 $(\mathcal{E}_{ extit{X}}^{ extit{y}}(e) \text{ replaces all occurrences of } x \text{ in } \mathcal{E}(e) \text{ by } y)$ 

#### **Example**

$$\begin{array}{lll} \mathcal{F}(\texttt{o.f} == \texttt{\lobel{loop}}(\texttt{o.f}) + \texttt{1}) &= \\ \mathcal{E}(\texttt{o.f}) = \mathcal{E}(\texttt{\lobel{loop}}(\texttt{o.f}) + \texttt{1}) &= \\ \mathcal{E}(\texttt{o.f}) = \mathcal{E}(\texttt{\lobel{loop}}(\texttt{o.f})) + \mathcal{E}(\texttt{1}) &= \\ \mathcal{E}(\texttt{o.f}) = \mathcal{E}^{\texttt{heap}}_{\texttt{heap}}(\texttt{o.f}) + \texttt{1} &= \\ \texttt{select}(\texttt{heap}, \texttt{o, f}) = \texttt{select}(\texttt{heapAtPre}, \texttt{o, f}) + \texttt{1} \end{array}$$

# Translation of Ensures and Signals Clauses

```
Given the normalised JML contract
/*@ public behavior
   0 . . .
   @ ensures E:
   @ signals (Throwable exc) S;
   @ ...
   0*/
Define
\mathcal{F}_{\mathsf{ensures}} = \mathcal{F}(\mathtt{E})
\mathcal{F}_{\mathsf{signals}} = \mathcal{F}(\mathtt{S})
```

```
Recall that S is either false, or it has the form

(exc instanceof ExcType1 ==> ExcPost1) && ...;
```

In the following, assume exc is fresh program variable of type Throwable

# Combining Signals and Ensures to post

The DL formula *post* is then defined as

$$(\texttt{exc} = \mathbf{null} \to \mathcal{F}_{\texttt{ensures}}) \; \& \; (\texttt{exc!} = \mathbf{null} \to \mathcal{F}_{\texttt{signals}})$$

#### Note:

Normalisation of normal\_behavior contract gives signals (Throwable exc) false; Then *post* is:

$$\begin{array}{ll} (\texttt{exc} = \texttt{null} \to \mathcal{F}_{\texttt{ensures}}) \ \& \ (\texttt{exc} \, ! = \texttt{null} \to \mathcal{F}_{\texttt{signals}}) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \to \mathcal{F}_{\texttt{ensures}}) \ \& \ (\texttt{exc} \, ! = \texttt{null} \to \mathcal{F}(\texttt{false})) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \to \mathcal{F}_{\texttt{ensures}}) \ \& \ (\texttt{exc} \, ! = \texttt{null} \to \texttt{false}) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \to \mathcal{F}_{\texttt{ensures}}) \ \& \ \texttt{exc} = \texttt{null} \\ \Leftrightarrow & \texttt{exc} = \texttt{null} \ \& \ \mathcal{F}_{\texttt{ensures}} \end{array}$$

#### Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- ▶ a precondition DL formula *pre* ✓,
- a postcondition DL formula post
- → a divergence indicator div ∈ {TOTAL, PARTIAL},
- a variant var a term of type any,
- ▶ a modifies set mod, either of type LocSet or \strictly\_nothing

SEFM: Proof Obligations

#### The Divergence Indicator

#### Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- ▶ a precondition DL formula pre
- a postcondition DL formula post
- ▶ a divergence indicator div ∈ {TOTAL, PARTIAL},
- a variant var a term of type any (postponed to later lecture),
- ▶ a modifies set mod, either of type LocSet or \strictly\_nothing

SEFM: Proof Obligations

# Translating Assignable Clauses: The DL Type LocSet

Assignable clauses are translated to

a term of type LocSet or the special value \strictly\_nothing

Intention: A term of type LocSet represents a set of locations

## **Definition (Locations)**

A location is a tuple (o, f) with  $o \in D^{\text{Object}}$ ,  $f \in D^{\text{Field}}$ 

Note: Location is a semantic and not a syntactic entity.

### The DL Type LocSet

```
Predefined type with D(LocSet) = 2^{Location}
and the functions (all with result type LocSet):
                                       empty set of locations: \mathcal{I}(empty) = \emptyset
 empty
                                       set of all locations, i.e., \mathcal{I}(\texttt{allLocs}) =
 allLocs
                                         \{(d, f)|f.a.\ d \in D^{\text{Object}}, f \in D^{\text{Field}}\}
 singleton(Object, Field)
                                       singleton set
 union(LocSet, LocSet)
 intersect(LocSet, LocSet)
 allFields(Object)
                                       set of all locations for the given object
 allObjects(Field)
                                       set of all locations for the given field;
                                       e.g., \{(d, f)|\text{f.a. }d \in D^{\text{Object}}\}
 arrayRange(Object, int, int)
                                       set representing all array locations in
                                       the specified range (both inclusive)
```

## Translating Assignable Clauses—Example

### **Example**

```
assignable \everything;
is translated into the DI term
```

allLocs

### **Example**

### Translating JML into Intermediate Format

#### Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- a precondition DL formula pre
- ▶ a postcondition DL formula post
- ▶ a divergence indicator  $div \in \{TOTAL, PARTIAL\}$   $\checkmark$ ,
- a variant var a term of type any (postponed),
- a modifies set mod, either of type LocSet or \strictly\_nothing



# From JML Contracts to Intermediate Format to Proof Obligations (PO)

```
public class A {
   /*@ public normal_behavior
   @ requires <Precondition>;
   @ ensures <Postcondition>;
   @ assignable <locations>;
   @*/
public int m(params) {..}
}
Intermediate Format

(pre, post, div, var, mod)

Translation

PO Generation

PO Generation

PO Generation
```

Proof obligation as DL formula

```
pre 
ightarrow \langle 	exttt{this.m(params);} 
angle \ (post \& frame)
```

# Generating a PO from the Intermediate Format: Idea

Given intermediate format of contract of m implemented in class C:



$$\mathit{pre} 
ightarrow \langle \mathtt{self.m(args)} \rangle (\mathit{post}$$
 &

frame )

correctness of assignable

(In case of div = PARTIAL, box modality is used)

# Generating a PO from Intermediate Format: Method Identification

$$pre \rightarrow \langle self.m(args) \rangle (post & frame)$$

- ▶ Dynamic dispatch: self.m(...) causes split into all possible implementations
- Special statement Method Body Statement:

Meaning: Placeholder for the method body of class C

# Generating a PO from Intermediate Format: Exceptions

$$pre \rightarrow \langle \texttt{self.m}(\texttt{args}) @C \rangle (post \& frame)$$

Postcondition post states either

- that no exception is thrown or
- ▶ that in case of an exception the exceptional postcondition holds

```
but: \langle \text{throw exc}; \rangle \varphi is trivially false

How to refer to an exception in post-state?

pre \rightarrow

exc = null;

try \{
self.m(args)@C
} catch (Throwable t) \{exc = t;\}

(post & frame)
```

(Recall: Normalistion and post-generation used program variable exc)

## The Generic Precondition genPre

```
pre \to \langle \texttt{exc=null}; \ try \ \{\texttt{self.m(args)@C}\} \ catch \ \dots \ \rangle (\textit{post \& frame}) is still not complete.
```

Additional properties (known to hold in Java, but not in DL), e.g.,

- ▶ this is not null
- created objects can only point to created objects (no dangling references)
- integer parameters have correct range

Need to make these assumption on initial state explicit in DL. Idea: Formalise assumption as additional precondition genPre

```
(genPre \land pre) \rightarrow \\ (exc=null; try {self.m(args)@C} catch ... )(post & frame)
```

## The Generic Precondition genPre

```
\begin{split} \textit{genPre} := & \quad \texttt{wellFormed(heap)} \\ & \quad \land \texttt{paramsInRange} \\ & \quad \land \texttt{self} \neq \textbf{null} \\ & \quad \land \texttt{boolean} :: \texttt{select(heap, self, < created>)} = \texttt{TRUE} \\ & \quad \land \texttt{C} :: \texttt{exactInstance(self)} \\ & \quad \land \texttt{exc} = \textbf{null} \end{split}
```

- wellFormed: predefined predicate; true iff. given heap is regular Java heap
- paramsInRange formula stating that the method arguments are in range
- ► C :: exactInstance: predefined predicate; true iff. given argument has C as exact type (i.e., is not of a subtype)

## The Generic Precondition genPre

```
(genPre \land pre) \rightarrow \\ \langle exc=null; try {self.m(args)@C} catch ... \rangle (post & frame) is still not complete.
```

▶ Need to refer to prestate in post, e.g. for old-expressions

```
(genPre \land pre) \rightarrow \{heapAtPre := heap\} 
\langle exc=null; try \{self.m(args)@C\} catch ... \rangle (post & frame)
```

(Reminder: heapAtPre was used in translation of \old in post)

# Generating a PO from Intermediate Format: Result Value

```
 \begin{split} (\textit{genPre} \land \textit{pre}) &\rightarrow \{\texttt{heapAtPre} := \texttt{heap}\} \\ & \langle \texttt{exc=null}; \ \texttt{try} \ \{\texttt{self.m(args)}\} \ \textbf{catch} \ \dots \ \rangle \\ & (\textit{post \& frame}) \end{split}  is still not complete.
```

► For non-void methods, need to refer to result in post

(Reminder: result was used in translation of \result in post)

# Generating a PO from Intermediate Format: The *frame* DL Formula

```
(genPre \land pre) \rightarrow \{\text{heapAtPre} := \text{heap}\}\ \langle \text{exc=null}; \text{ try } \{\text{self.m(args)}\} \text{ catch } \dots \rangle (post \& frame)

If mod = \text{strictly\_nothing} \text{ then } frame \text{ is defined}

\forall o; \forall f; (\text{select(heapAtPre}, o, f) = \text{select(heap}, o, f))
```

# Generating a PO from Intermediate Format: The *frame* DL Formula

If mod is a location set, then frame is defined as:

```
\forall o; \forall f; ( select(heaptAtPre, o, < created>) = FALSE \\ \lor select(heapAtPre, o, f) = select(heap, o, f) \\ \lor (o, f) \in \{heap := heapAtPre\} mod)
```

States that any location (o, f)

- belongs to an object was not (yet) created before the method invocation, or
- holds the same value after the invocation as before the invocation, or
- belongs to the modifies set (evaluated in the pre-state).

## **Examples**

Demo

### Literature for this Lecture

### Essential

KeY Quicktour see course page, under 'Links, Papers, and Software'

SEFM: Proof Obligations CHALMERS/GU 161014 49 / 49