## Formal Methods for Software Development Proof Obligations

Wolfgang Ahrendt

13 October 2017

## **This Part**

#### making the connection between

JML

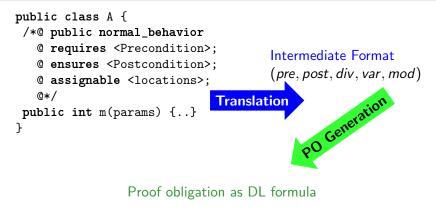
and

Dynamic Logic / KeY

- generating,
- understanding,
- and proving

DL proof obligations from JML specifications

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



```
pre 
ightarrow \langle \texttt{this.m(params);} 
angle \ (post \ \land \ 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

## New JML Feature: Nested Specification Cases

```
method charge() has nested specification case:
@ public normal behavior
@ requires amount > 0;
0 {
0
    requires amount + balance < limit && isValid()==true;
    ensures \result == true:
0
    ensures balance == amount + \old(balance);
0
0
    assignable balance;
0
0
    also
0
0
    requires amount + balance >= limit;
0
    ensures \result == false;
0
    ensures unsuccessfulOperations
0
            == \old(unsuccessfulOperations) + 1;
0
    assignable unsuccessfulOperations;
```

FMSD: Proof Obligations

nested specification cases allow to factor out common preconditions

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

```
(previous page) ... expands to
@ public normal_behavior
@ requires R;
@ requires R1;
@ ensures E1;
@ assignable A1;
0
 also
0
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
0
    ensures \result == true;
0
    ensures balance == amount + \old(balance);
0
    assignable balance;
0
0
    also
0
0
    requires amount + balance >= limit;
0
    ensures \result == false;
0
    ensures unsuccessfulOperations
             == \old(unsuccessfulOperations) + 1;
0
0
    assignable unsuccessfulOperations;
0
  1}
expands to ... (next page)
FMSD: Proof Obligations
                               CHALMERS/GU
                                                         171013
                                                               8 / 44
```

```
(previous page) ... expands to
@ public normal_behavior
@ requires amount > 0;
@ requires amount + balance < limit && isValid()==true;</pre>
@ 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;
FMSD: Proof Obligations
                              CHALMERS/GU
                                                        171013
```

9 / 44

# Normalisation: Making Implicit Information Explicit

#### Implicit Information

- Meaning of normal\_ and exceptional\_behavior
- non\_null by default
- \invariant\_for(this) in requires, ensures, signals clauses

#### Turn into general behavior spec. case

- 1. Add to
  - > normal\_behavior the clause signals (Throwable t) false;
  - exceptional\_behavior the clause ensures false;
- 2. Replace normal\_behavior/exceptional\_behavior by behavior

## Normalisation: Making Implicit Information Explicit

### **Implicit Information**

- Meaning of normal\_ and exceptional\_behavior
- non\_null by default
- \invariant\_for(this) in requires, ensures, signals clauses

#### Making non\_null explicit in method specifications

- Where nullable is absent, add o != null to preconditions (for parameters<sup>a</sup>) and postconditions (for return values<sup>a</sup>).
   E.g., for method void m(Object o) add requires o != null;
- Thereafter add nullable, where absent, to all parameter<sup>a</sup> and return type<sup>a</sup> declarations

<sup>&</sup>lt;sup>a</sup>of reference type

## Normalisation: Making Implicit Information Explicit

### Implicit Information

- Meaning of normal\_ and exceptional\_behavior
- non\_null by default
- \invariant\_for(this) in requires, ensures, signals clauses

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

- 1. Add explicit \invariant\_for(this) to non-helper method specs:
  - requires \invariant\_for(this);
  - ensures \invariant\_for(this);
  - signals (Throwable t) \invariant\_for(this);
- 2. Thereafter add helper, where absent, to all methods

## Normalisation: Example

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

becomes

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

## Normalisation: Example

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

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

## Normalisation: Example

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

## Next Normalisation Steps (Not detailed)

- Expanding pure modifier:
  - add to each specification case
    - assignable \nothing;
    - diverges false;
  - remove pure

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

## Normalisation: Clause Contraction

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

For instance,

```
/*@ public behavior
@ requires R1;
@ requires R2;
@ ensures E1;
@ ensures E2;
@ signals (T1 exc) S1;
@ signals (T2 exc) S2:
@*/
```

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

## 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\},\$
- a variant var a term of type any
- a modifies set mod, either of type LocSet or \strictly\_nothing

# 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"

# 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), pretty-printed as 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
- true, false are formulas,

boolean constants TRUE, FALSE are terms

Atomic formulas take terms as arguments; e.g.:

```
x - y < 5</li>
b = TRUE
```

## Translating Boolean JML Expressions

$$\begin{array}{rcl} \mathcal{F}(v) & = & v = \text{TRUE} \\ \mathcal{F}(o.f) & = & \mathcal{E}(o.f) = \text{TRUE} \\ \mathcal{F}(m()) & = & \mathcal{E}(m)() = \text{TRUE} \\ \mathcal{F}(!b_{-}0) & = & !\mathcal{F}(b_{-}0) \\ \mathcal{F}(b_{-}0 & \& \& b_{-}1) & = & \mathcal{F}(b_{-}0) & \& \mathcal{F}(b_{-}1) \\ \mathcal{F}(b_{-}0 & = & b_{-}1) & = & \mathcal{F}(b_{-}0) & | \mathcal{F}(b_{-}1) \\ \mathcal{F}(b_{-}0 & = & b_{-}1) & = & \mathcal{F}(b_{-}0) & | \mathcal{F}(b_{-}1) \\ \mathcal{F}(b_{-}0 & = & b_{-}1) & = & \mathcal{F}(b_{-}0) & | \mathcal{F}(b_{-}1) \\ \mathcal{F}(b_{-}0 & = & b_{-}1) & = & \mathcal{F}(b_{-}0) & | \mathcal{F}(b_{-}1) \\ \mathcal{F}(b_{-}0 & = & e_{-}1) & = & \mathcal{E}(e_{-}0) & = \mathcal{E}(e_{-}1) \\ \mathcal{F}(e_{-}0 & = & e_{-}1) & = & \mathcal{E}(e_{-}0) & | \mathcal{E}(e_{-}1) \\ \mathcal{F}(e_{-}0 & = & e_{-}1) & = & \mathcal{E}(e_{-}0) & | \mathcal{E}(e_{-}1) \\ \mathcal{F}(e_{-}0 & = & e_{-}1) & = & \mathcal{E}(e_{-}0) & | \mathcal{E}(e_{-}1) \\ \end{array}$$

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 Quantified formulas over reference types:

```
 \mathcal{F}((\langle \mathbf{forall } T x; e_0; e_1)) = \\ \langle \mathbf{forall } T x; ( \\ (!x=\mathbf{null} \& x. < created > = TRUE \& \mathcal{F}(e_0)) \\ - > \mathcal{F}(e_1)) \\ \mathcal{F}((\langle \mathbf{exists } T x; e_0; e_1)) = \\ \langle \mathbf{exists } T x; ( \\ (!x=\mathbf{null} \& x. < created > = TRUE \& \mathcal{F}(e_0)) \\ \& \mathcal{F}(e_1)) \\ \end{cases}
```

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

$$\mathcal{F}((\texttt{forall int } x; e_0; e_1)) = \\ \texttt{forall int } x; ((\texttt{inInt}(x) \& \mathcal{F}(e_0)) \rightarrow \mathcal{F}(e_1))$$

inInt (similar inLong, inByte):
 Predefined predicate symbol with fixed interpretation
 Meaning: Argument is within the range of the Java int datatype.

 $\mathcal{F}(\texttt{invariant_for}(e)) = \text{Object ::<inv>(heap, \mathcal{E}(e))}$ 

- \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 (during proof construction) 'Object ::<inv>(heap, o)' to the invariant of T
- Dbject ::<inv>(heap, o) pretty printed as o.<inv>
- Read 'invariant of o'

## Translating JML into Intermediate Format

Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

- a precondition DL formula pre V,
- ► 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

## **Translation of Ensures Clauses**

What is missing for ensures clauses?

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

#### Translating \result

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

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

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

# Translating \old Expressions

 $\label{eq:local} \$  evaluates e in the prestate of the method Accesses to heap must be evaluated w.r.t. to the 'old' heap

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

$$\mathcal{E}(\texttt{\old}(e)) = \mathcal{E}_{\texttt{heap}}^{\texttt{heapAtPre}}(e)$$

 $(\mathcal{E}_x^y(e) \text{ replaces all occurrences of } x \text{ in } \mathcal{E}(e) \text{ by } y)$ 

## Example

## **Translation of Ensures and Signals Clauses**

Given the normalised JML contract

```
/*@ public behavior
@ ...
@ ensures E;
@ signals (Throwable exc) S;
@ ...
@*/
```

```
Define \mathcal{F}_{ensures} = \mathcal{F}(E)
```

```
\mathcal{F}_{\mathsf{signals}} = \mathcal{F}(\mathtt{S})
```

```
Recall (p.16) 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 Ensures and Signals to** post

The DL formula *post* is then defined as

$$(\texttt{exc} = \texttt{null} \rightarrow \mathcal{F}_{\texttt{ensures}}) \ \land \ (\texttt{exc} \neq \texttt{null} \rightarrow \mathcal{F}_{\texttt{signals}})$$

#### Important special case:

Normalisation of normal\_behavior contract gives

signals (Throwable exc) false;

In that case, *post* is:

$$\begin{array}{ll} (\texttt{exc} = \texttt{null} \rightarrow \mathcal{F}_{\texttt{ensures}}) & \land & (\texttt{exc} \neq \texttt{null} \rightarrow \mathcal{F}_{\texttt{signals}}) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \rightarrow \mathcal{F}_{\texttt{ensures}}) \land & (\texttt{exc} \neq \texttt{null} \rightarrow \mathcal{F}(\texttt{false})) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \rightarrow \mathcal{F}_{\texttt{ensures}}) \land & (\texttt{exc} \neq \texttt{null} \rightarrow \texttt{false}) \\ \Leftrightarrow & (\texttt{exc} = \texttt{null} \rightarrow \mathcal{F}_{\texttt{ensures}}) \land & \texttt{exc} = \texttt{null} \\ \Leftrightarrow & \texttt{exc} = \texttt{null} \land \mathcal{F}_{\texttt{ensures}} \end{array}$$

## Translating JML into Intermediate Format

## Intermediate format for contract of method m

(pre, post, div, var, mod)

with

- a precondition DL formula pre V,
- a postcondition DL formula post V,
- ▶ 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

## The Divergence Indicator

div =
{ TOTAL if normalised JML contract contains clause diverges false;
 PARTIAL if normalised JML contract contains clause diverges true;

# 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  $o \in D^{\texttt{Object}}$ ,  $f \in D^{\texttt{Field}}$ 

## The DL Type LocSet

Predefined type with  $D(LocSet) = 2^{Location}$ and the functions (all with result type LocSet):

empty allLocs

singleton(Object,Field)
union(LocSet,LocSet)
intersect(LocSet,LocSet)
allFields(Object)
allObjects(Field)

arrayRange(Object, int, int)

empty set of locations:  $\mathcal{I}(\text{empty}) = \emptyset$ set of all locations, i.e.,  $\mathcal{I}(\texttt{allLocs}) = \{(d, f) | f.a. \ d \in D^{\texttt{Object}}, f \in D^{\texttt{Field}}\}$ singleton set

set of all locations for the given object set of all locations for the given field; e.g.,  $\{(d, f) | f.a. \ d \in D^{\text{Object}}\}$ set representing all array locations in the specified range (both inclusive)

## Translating Assignable Clauses—Example

## Example

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

allLocs

## Example

assignable this.next, this.content[5..9];

is translated into the DL term

union(singleton(self,next), arrayRange(self.content,5,9)

## Translating JML into Intermediate Format

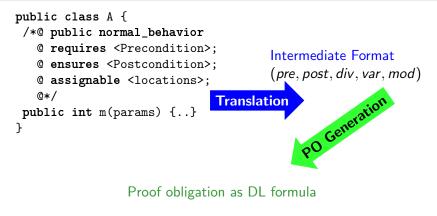
Intermediate format for contract of method m

(pre, post, div, var, mod)

#### with

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

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



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

# Generating a PO from the Intermediate Format: Idea

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

```
(pre, post, TOTAL, var, mod)
                             PO
                            Generation
pre \rightarrow (self.m(args))(post \land
                                               trame
                                           correctness of
                                             assignable
```

In case of div = PARTIAL, box modality is used

## Generating a PO from Intermediate Format: Method Identification

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

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

#### m(args)@C

Meaning: implementation of m in class C

# Generating a PO from Intermediate Format: Exceptions

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

Postcondition post states either

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

but:  $\langle \mathbf{throw} \; \mathsf{exc}; \rangle \varphi$  is trivially false

How to refer to an exception in post-state?

```
pre \rightarrow
exc = null;
\left\langle \begin{array}{c} try \{ \\ self.m(args)@C \\ \} catch (Throwable e){exc = e;} \end{array} \right\rangle (post \land frame)
```

Recall: generation of *post* (p.28) uses program variable exc

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

## The Generic Precondition genPre (background info)

- wellFormed(h): predefined predicate; true iff. h is regular Java heap
- C :: exactInstance(o): predefined predicate; true iff. o has exact type C (not just subtype of C)
- paramsInRange formula stating that method arguments are in range

## The Generic Precondition genPre

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

is still not complete.

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

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

Recall: heapAtPre was used in translation of \old, p.26

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

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

If  $mod = \texttt{strictly_nothing}$  then *frame* is defined as:

$$\forall o; \forall f; (o.f = o.f@\texttt{heapAtPre})$$

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

$$\begin{array}{l} (\textit{genPre} \land \textit{pre}) \rightarrow \{\texttt{heapAtPre} := \texttt{heap}\} \\ \quad & \langle \texttt{exc=null}; \ \texttt{try} \ \{\texttt{self.m(args)}\} \ \texttt{catch} \ \dots \ \rangle \\ \quad & (\textit{post} \ \land \ \textit{frame}) \end{array}$$

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

$$\forall o; \forall f; ((o, f) \in \{\text{heap} := \text{heapAtPre}\} mod \ \lor o. < \text{created} > @ \text{heaptAtPre} = FALSE \ \lor o.f = o.f @ \text{heapAtPre})$$

Says that every location (o, f) either

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

# Generating a PO from Intermediate Format: Result Value

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

is still not complete.

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

$$\begin{array}{l} (genPre \land pre) \rightarrow \{\texttt{heapAtPre} := \texttt{heap}\} \\ \land \texttt{(exc=null; try {result = self.m(args)} catch ... } \\ (post \land frame) \end{array}$$

Recall: \result was translated to program variable result, see p.25

## **Examples**

## Demo