VCC Syntax Refresh

Ernie Cohen, Michał Moskal, Wolfram Schulte, Stephan Tobies

This document describes our initial thoughts about updating the annotation language of VCC to make it simpler and more consistent. Our starting point is the existing ACLS specification (ANSI/ISO C Specification Language), which is used by the Frama-C project and is available from

<http://frama-c.cea.fr/download/acsl_1.4.pdf>

This document follows the structure of the ACLS spec and outlines the major differences and how to plan (if any) to resolve them with the ACSL standard.

### 1.2 Generalities

Instead of using comments to hide the annotations from the normal C compiler, we will use a macro:

#ifdef VERIFY  
#define spec \_\_vcc\_spec  
#else  
#define spec(…)  
#endif

This allows preprocessing to work as usual. Multi-line annotations are allowed and need to be separated by semicolon:

spec( requires x < 5;  
 ensures y > 10 )

Users of VCC are free to #define a synonym for spec and, e.g., use ‘\_’ to reduce the visual clutter.

### 2.1 Lexical rules

As ACSL, we will support the use of a number of Unicode chars in annotations (e.g. ≤, ∀, ∧, ⇒); we might want to consider VS editor support to replace \forall etc. with these special forms using an ‘autocorrect’ approach.

Identifiers can start with the backslash character and any annotation-specific keyword does so (see ‘\forall’ below), except when the keyword is at the start of a spec block (like the ‘requires’ in the example above) or after a semicolon that starts new annotation (like the ‘ensures’ above).

### 2.2 Logic expressions

* We do **not** distinguish between terms and predicates. We follow the ACSL syntax for quantification:

\forall int x,y, uint z; foo(x,y,z)

* We allow triggers after the ‘;’ as in \forall obj x; {bar(x)} foo(bar(x))
* We add the support for \let .
* For functional modifiers of struct or maps, we add the syntax { s \with .id = v } in addition to the use of ‘for’ from ACSL; this replaces our current syntax s / { .id = v }.
* We do not add the bitwise Boolean operators ‘-->’ or ‘<-->’.
* We do not support consecutive comparison operators of the form a <= b < c because it makes moving code between implementation and spec harder.

### 2.2.4 Integer arithmetic and machine integers

ACLS diverges significantly from VCC in its use of mathematical integers. In ACLS specification, all machine integers types are implicitly converted to the mathematical integer type ‘integer’ (‘mathint’ in VCC, which we will rename). Conversion back to machine integer types must be explicit and will always succeed by taking the “unique value of the corresponding type that is congruent to the mathematical result modulo the cardinal of this type, that is ”

We are currently considering this approach, too. It will solve a number of semantic problems that we are encountering, with the following additions:

* Using the notation \C(e), the expression e will be evaluated using (unchecked) C semantics.
* We want to have a way to express checked semantics that assert absence of overflows in implementation code. How?

We would be really interested in getting feedback on this point!

### 2.2.5 Real numbers and floating point numbers

For now, no support beyond what we currently have.

### 2.2.6 Structures, unions, arrays in logic

ACSL uses a separate syntax for logic type definitions. C already has syntax for typedefs and would use that instead of a second mechanism for defining types.

We will introduce a separate \record structured type, which replaces our current kludge of ‘struct vcc(record) { … }’

Does ACSL have a map type? What about lambda expressions?

### 2.2.8 Address operator

As our ownership system relies on pointers also for spec objects, we do support taking of addresses of spec objects and stick with our current distinction between implementation pointers (T\*) and specification pointers (T^).

#### 2.3 Function Contracts

* Function contracts should follow the function prototype, not precede it.
* The ACSL annotations for termination seem to make sense but are orthogonal to our current efforts and can optionally be added in later. If so, support general termination measures, not only integers.
* Named behaviors have low priority, but if we should adopt them, we would follow their example
* In ACSL, no writes clause means unspecified with a provision for tool inference should the function be defined. For use, no writes clause means no writes, and we tend to like to keep it that way.
* ACSL allows merging of multiple contracts. We think that having all contracts in one place is better, modulo a support for public/private contracts, which have been requested for VCC, but for which we have not proper answer yet.
* For framing, we would diverge from ACSL and write ‘writes p’ where they would write ‘assigns \*p’; we think that the pointer-oriented approach is more suitable for our object-oriented memory model. We would use ‘writes’ instead of ‘assigns’ to disambiguate our approach from theirs.
* We will adopt the following syntax for arrays:
  + as\_array(p, 10)will be replaced by (int[10])p
  + array\_range(p, 10) will be replaced by p + (0 .. 9) or, equivalently, &p[0 .. 9] I am still not 100% happy with this second thing.

### 2.4 Statement annotations

* Asserts become their own annotation elements: spec( assert p ) instead of assert(p) so we do not collide with assert.h; if desired, the assert function from assert.h can get a precondition that requires the condition to hold.
* Add an assume statement
* Similar to function contracts, loop contracts should follow the loop head and not precede it.
* It seems that we already implement the same notion of loop invariant as they do, no semantic difference there.
* We are unclear about the use of invariants in the middle of a loop body – we will need to check how this would translate to Boogie.
* Their general invariants also would provide support for loops created by gotos – we should consider adding support for that, too.
* ACSL’s approach of using \at with labels seems questionable: what if the label had not been passed on the way to the current state, or multiple times? Instead, we stick with our support of states as first-class citizens. We will need clean up our story regarding the use of old for locals and non-head parameters; also, the different semantics of \old in loops needs to be re-considered, I think.

### 2.6 Logical specification

We would change our current way of specifying pure functions to:

spec(   
 pure bool isPositive(int x)  
 ensures \result == x > 0;  
)

For the moment, such pure functions must not have a body. In the future, we may also consider allowing pure functions with bodies, which then provide a witness. For compatibility with ACSL, we could consider adding their syntax, too, but would prefer ours because we do not need an entirely new class of declarations. Are we limiting this to the from \result == … ?

We also want to support spec functions with bodies, which would be indicated by the use of the keyword ‘procedure’ as in:

spec(  
 procedure bool update(obj o)  
 { … }  
)

Suggestions for a better keyword than ‘procedure’ are welcome!

Their use of lemmas is not suitable in our context, though we need ways of specifying axioms (which we should syntactically match to the rest of our syntax) and a way to achieve what we currently do with bv\_lemma. Abstractly, this is nothing but a hint how certain assertions should be proven before they can be used, which would be reflected in the following sytax:

{  
 …  
 spec( assert {bv} \forall int i; … )  
}

Here, the {bv} serves as a pragma to VCC that suggest how the assert should be proved. One could think of other pragmas for similar purposes, thus we support them in this more general fashion.