První kroky s Frama-C

Jan Tušil

16. 3. 2015

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

Závěr

Výhody Omezení

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

7ávěr

Výhody

Materiály k prezentaci

▶ https://github.com/h0nzZik/Frama-C_Examples

Jak získat?

- ▶ Balík frama-c v repozitářích Fedory i Debianu
- Alt-Ergo (theorem prover) tamtéž (hodí se pro WP plugin)



Modulární platforma pro statickou analýzu C (ISO C99) kódu



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro

Pluginy





- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu

Pluginy



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session

Pluginy



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
- Pluginy



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
 - Manipulace s AST
- Pluginy



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
 - Manipulace s AST
- Pluginy
 - Analýza závislostí



- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
 - Manipulace s AST
- Pluginy
 - Analýza závislostí
 - Abstraktní interpretace





- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
 - Manipulace s AST
- Pluginy
 - Analýza závislostí
 - Abstraktní interpretace
 - Deduktivní verifikace





- Modulární platforma pro statickou analýzu C (ISO C99) kódu
- Jádro
 - Parsování, preprocessing, "linkování" a normalizace kódu
 - Ukládání/načítání session
 - Uchovávání výsledků práce pluginů
 - Manipulace s AST
- Pluginy
 - Analýza závislostí
 - Abstraktní interpretace
 - Deduktivní verifikace
- Některé pluginy umožňují ověřování programu vůči specifikaci



Možné specifikace:

- Možné specifikace:
 - Za běhu programu nenastane chyba

- Možné specifikace:
 - Za běhu programu nenastane chyba (?)

- Možné specifikace:
 - ► Za běhu programu nenastane chyba (?)
 - Chování programu je definované (dle standardu jazyka).

- Možné specifikace:
 - Za běhu programu nenastane chyba (?)
 - Chování programu je definované (dle standardu jazyka).
 - ► Funkce nemodifikuje globální proměnné

- Možné specifikace:
 - Za běhu programu nenastane chyba (?)
 - Chování programu je definované (dle standardu jazyka).
 - Funkce nemodifikuje globální proměnné
 - Funkce zachovává vlastnost používané datové struktury

- Možné specifikace:
 - Za běhu programu nenastane chyba (?)
 - Chování programu je definované (dle standardu jazyka).
 - Funkce nemodifikuje globální proměnné
 - Funkce zachovává vlastnost používané datové struktury
- Potřeba jazyka pro zápis specifikace.

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

7ávěr

Výhody Omezen

Osnova

Úvod

Jazyk ACSL

Obecné informace

Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE)

7ávěr

Výhody

ACSL - Co je to za jazyk?

- ► ANSI / ISO C Specification language
- Vznikl pro potřeby Frama-C
- Vestavěný v komentářích C kódu.
- Asserty
- Invarianty
- Funkční kontrakty (DbC)
- http://frama-c.com/acsl.html

ACSL - vyjadřovací schopnosti

- C operátory a datové typy
- Matematické datové typy
- Prvořádová logika (s rozšířeními)

Osnova

Úvod

Jazyk ACSL

Obecné informace

Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE)

7ávěr

Výhody

Asserty

```
int x = 17;
/*@ assert x > 5 */
```

- Základní specifikační jednotka
- Tvrzení o stavu programu v daném bodě.
- (ACSL specifikace zabudována v komentáři)

Kvantifikátory

```
int array[4] = {-15, 3, 17, 104};
/*@
assert \forall integer i, j;
0 <= i <= j < 4 ==> array[i] <= array[j];
*/</pre>
```

- Vázaná proměnná má daný typ
- Typ může být uživatelem definovaný (typedef, struct)
- ▶ integer označuje (matematické) celé číslo

```
int array[4] = {-15, 3, 17, 104};
/*@ assert \valid(array + (0..3)); */
```

- Ukazuje pointer na platnou paměť?
- Predikát \valid se vyhodnotí na true, pokud je každý pointer z dané množiny platný.

```
int array[4] = {-15, 3, 17, 104};
/*@ assert \valid(array + (0..3)); */
```

- Ukazuje pointer na platnou paměť?
- Predikát \valid se vyhodnotí na true, pokud je každý pointer z dané množiny platný.
- ightharpoonup Zápis 0..3 označuje range (tedy množinu $\{0,1,2,3\}$)

```
int array[4] = {-15, 3, 17, 104};
/*@ assert \valid(array + (0..3)); */
```

- Ukazuje pointer na platnou paměť?
- Predikát \valid se vyhodnotí na true, pokud je každý pointer z dané množiny platný.
- ► Zápis 0..3 označuje range (tedy množinu {0,1,2,3})
- Range je možné přičítat k ostatním hodnotám

```
int array[4] = {-15, 3, 17, 104};
/*@ assert \valid(array + (0..3)); */
```

- Ukazuje pointer na platnou paměť?
- Predikát \valid se vyhodnotí na true, pokud je každý pointer z dané množiny platný.
- ► Zápis 0..3 označuje range (tedy množinu {0,1,2,3})
- Range je možné přičítat k ostatním hodnotám
- Význam: array ukazuje na platnou paměť o délce alespoň 4 intů.

```
int array[4] = {-15, 3, 17, 104};
/*@ assert \valid(array + (0..3)); */
```

- Ukazuje pointer na platnou paměť?
- Predikát \valid se vyhodnotí na true, pokud je každý pointer z dané množiny platný.
- ► Zápis 0..3 označuje range (tedy množinu {0,1,2,3})
- Range je možné přičítat k ostatním hodnotám
- Význam: array ukazuje na platnou paměť o délce alespoň 4 intů.
- Nástroj pracuje s obvyklou pointer aritmetikou



Uživatelské predikáty

```
/*@
predicate is_sorted ( int *array, integer len ) =
\forall integer i, j; 0 <= i <= j < len
==> array[i] <= array[j];
*/

    Predikát lze využít později
/*@ assert is_sorted ( array, 4 ); */</pre>
```

Invarianty smyček

```
int arr[7];
[ ... ]
/*@ loop invariant is_sorted(arr, i) */
for ( int i = 0; i < 7; i++ ) {
        [ ... ]
}</pre>
```

Invarianty smyček

```
int arr[7];
[ ... ]
/*@ loop invariant is_sorted(arr, i) */
for ( int i = 0; i < 7; i++ ) {
        [ ... ]
}</pre>
```

Invariant musí být platný při (prvním) vstupu do smyčky

Invarianty smyček

```
int arr[7];
[ ... ]
/*@ loop invariant is_sorted(arr, i) */
for ( int i = 0; i < 7; i++ ) {
        [ ... ]
}</pre>
```

- Invariant musí být platný při (prvním) vstupu do smyčky
- Každý průchod smyčkou musí invariant zachovat

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
  */
void sort ( int *array, size_t len);
```

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
 */
void sort ( int *array, size_t len);
```

Co funkce požaduje?

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
 */
void sort ( int *array, size_t len);
```

- Co funkce požaduje?
- Co funkce garantuje?

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
 */
void sort ( int *array, size_t len);
```

- Co funkce požaduje?
- Co funkce garantuje?
- Co funkce zachovává / modifikuje?

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
 */
void sort ( int *array, size_t len);
```

- Co funkce požaduje?
- Co funkce garantuje?
- Co funkce zachovává / modifikuje?
- Paradigma "Design by Contract"

```
/*@
requires \valid ( array + (0 .. ( len-1 ) ) );
ensures is_sorted ( array, len );
assigns \nothing;
  */
void sort ( int *array, size_t len);
```

- Co funkce požaduje?
- Co funkce garantuje?
- Co funkce zachovává / modifikuje?
- Paradigma "Design by Contract"
- Kde je problém?

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

7ávěr

Výhody Omezen

\$ frama-c hello.c -val

- \$ frama-c hello.c -val
- Spustí Value plugin analýzu možných hodnot proměnných.

- \$ frama-c hello.c -val
- Spustí Value plugin analýzu možných hodnot proměnných.
- Jak ověřit větší projekt z více souborů?

- \$ frama-c hello.c -val
- Spustí Value plugin analýzu možných hodnot proměnných.
- Jak ověřit větší projekt z více souborů?
- Všechny je vypíšeme na řádce.

- \$ frama-c hello.c -val
- Spustí Value plugin analýzu možných hodnot proměnných.
- Jak ověřit větší projekt z více souborů?
- Všechny je vypíšeme na řádce.
- Soubor hello.c je v adresáři 01_hello

Makra a náhrady textu

- Makra a náhrady textu
- Vkládání (hlavičkových) souborů

- Makra a náhrady textu
- Vkládání (hlavičkových) souborů
- Možno specifikovat, v jakých adresářích se hlavičky nacházejí

- ▶ Výchozí: gcc -C -E -I.
- ► Možno předefinovat přepínačem -cpp-command

- ▶ Výchozí: gcc -C -E -I.
- Možno předefinovat přepínačem -cpp-command
- Frama-C dodává oanotovanou stadardní knihovnu (libc)

- ▶ Výchozí: gcc -C -E -I.
- Možno předefinovat přepínačem -cpp-command
- Frama-C dodává oanotovanou stadardní knihovnu (libc)
- Frama-C umí předzpracovat i anotace (s GCC)

- ▶ Výchozí: gcc -C -E -I.
- Možno předefinovat přepínačem -cpp-command
- Frama-C dodává oanotovanou stadardní knihovnu (libc)
- Frama-C umí předzpracovat i anotace (s GCC)
- Rozpracovaný projekt je možné uložit a znovu načíst

- ▶ Výchozí: gcc -C -E -I.
- Možno předefinovat přepínačem -cpp-command
- Frama-C dodává oanotovanou stadardní knihovnu (libc)
- Frama-C umí předzpracovat i anotace (s GCC)
- Rozpracovaný projekt je možné uložit a znovu načíst
- frama-c-gui -load frama_project

- ▶ Výchozí: gcc -C -E -I.
- Možno předefinovat přepínačem -cpp-command
- Frama-C dodává oanotovanou stadardní knihovnu (libc)
- Frama-C umí předzpracovat i anotace (s GCC)
- Rozpracovaný projekt je možné uložit a znovu načíst
- frama-c-gui -load frama_project
- Viz soubor build.mk

Drobnosti

- ▶ RTE plugin generuje anotace pro obvyklé runtime chyby
- ► Kombinace RTE + Value může prokázat absenci runtime chyb
- Uložený projekt je možné načíst do programu frama-c-gui

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů

Value plugin (+ RTE)

WP plugin

7ávěr

Výhody

Omezen

Value plugin - Principy

- Abstraktní interpretace
- Počítá variační domény proměnných
- Overaproximace dokazuje korektnost

Variační domény

- Množina možných hodnot, které může obsahovat daná proměnná.
- Různé způsoby zápisu
 - ▶ Výčtem {2, 12, 22, 32, 42}
 - ▶ Intervalem [2...42], 2%10

Variační domény

- Množina možných hodnot, které může obsahovat daná proměnná.
- Různé způsoby zápisu
 - ▶ Výčtem {2, 12, 22, 32, 42}
 - ▶ Intervalem [2...42], 2%10 (stejná množina)

Odbočka - RTE plugin

RunTime Error annotation generator

Odbočka - RTE plugin

- RunTime Error annotation generator
- Generuje asserty pro obvyklé runtime chyby a nespecifikované chování

Odbočka - RTE plugin

- RunTime Error annotation generator
- Generuje asserty pro obvyklé runtime chyby a nespecifikované chování
- Je využíván složitějšími pluginy

Ukázka (01_hello/hello.c) - 1

```
int main(int argc, char **argv)
   int retres:
   int array[25]:
   int idx:
 /*@ assert rte: index_bound: 0 ≤ argc; */
/*@ assert rte: index bound: argc < 25: */</p>
   array[argc] = 7;
idx = Frama C interval(-1,26);
   idx /= 2;
/*@ assert rte: index_bound: 0 ≤ idx; */
/*@ assert rte: index bound: idx < 25: */</p>
   array[idx] = 0x1234;
/*@ assert rte: signed overflow: -2147483648 ≤ idx*2; */
/*@ assert rte: signed overflow: idx*2 ≤ 2147483647; */
   idx = idx * 2 + 3:
/*@ assert rte: index bound: 0 ≤ idx; */
/*@ assert rte: index bound: idx < 25; */</p>
   array[idx] = 0x4567;
/*@ assert rte: signed overflow: idx+28 ≤ 2147483647; */
   idx += 28:
/*@ assert rte: index bound: 0 ≤ idx; */
/*@ assert rte: index bound: idx < 25; */</p>
   arrav[idx] = 15:
    retres = 0:
    return retres;
```

Ukázka (01_hello/hello.c) - 2

```
$ git clone https://github.com/h0nzZik/Frama-C_Examples.git
$ cd Frama-C_Examples/01_hello
$ make
```

- \$ frama-c-gui -load project_after_analysis
 - Informace vypsané během analýzy jsou k dispozici i v gui
 - Zajímavé jsou řádky začínající hello.c:123: [value]
 - Gui zobrazuje variační domény hodnoty proměnných na kartě "Information"

Ukázka (01_hello/hello.c) - 3

```
int main ( int argc, char *argv[] ) {
/*@ assert rte: index_bound: 0 <= argc; */
/*@ assert rte: index_bound: argc < 25; */
array[argc] = 7;</pre>
```

hello.c:17:[value] Assertion 'rte,index_bound' got status unknown.

- Nelze ověřit, že zápis do pole proběhne v pořádku.
- Mimochodem, tyto asserty v původním zdrojovém kódu nejsou, byly vygenerovány pluginem RTE. Proto jsou označeny identifikátorem rte.

```
idx = Frama_C_interval(-1,26);
idx /= 2;
/*@ assert rte: index_bound: 0 < idx; */
/*@ assert rte: index_bound: idx < 25; */
array[idx] = 0x1234;
hello.c:20:[value] Assertion 'rte,index_bound' got
status valid.</pre>
```

- ► Funkce Frama_C_interval vrací nedeterministickou hodnotu ze zadaného intervalu. Je deklarována (i s anotacemi) v souboru FRAMA_C_SHARE/builtin.h.
- Value plugin spočítá, že po dělení bude proměnná idx ležet v intervalu [0...13]

```
/*@ assert rte: signed_overflow: (int)(idx*2)+3 <= 214748364
/*@ assert rte: signed_overflow: -2147483648 <= idx*2; */
/*@ assert rte: signed_overflow: idx*2 <= 2147483647; */
idx = idx * 2 + 3;</pre>
```

- Mnoho lidí neví, že výsledek znaménkového přetečení v jazyce
 C není definován. RTE plugin to ale ví.
- ▶ Value určil $idx \in [3..29], 1\%2$

```
/*@ assert rte: index_bound: 0 <= idx; */
/*@ assert rte: index_bound: idx < 25; */
array[idx] = 0x4567;</pre>
```

- Value plugin ví, že jistě idx ∈ [3..29], 1%2, tedy první assert projde.
- Ale protože Value pracuje s overaproximací, nedokáže rozhodnout, zda dojde k narušení druhého assertu.
- Pokud dojde k runtime chybě, nemá smysl pokračovat v běhu programu. Z toho důvodu po provedení uvedeného příkazu Value plugin uvažuje pouze hodnoy z intervalu [3..23], 1%2.

```
idx += 28;
/*@ assert rte: index_bound: 0 <= idx; */
/*@ assert rte: index_bound: idx < 25; */
array[idx] = 15;</pre>
```

hello.c:26:[value] Assertion 'rte,index_bound' got status invalid (stopping propagation).

Po přičtení bude hodnota proměnné idx větší než 25. Assert tedy neprošel.

```
idx += 28;
/*@ assert rte: index_bound: 0 <= idx; */
/*@ assert rte: index_bound: idx < 25; */
array[idx] = 15;</pre>
```

hello.c:26:[value] Assertion 'rte,index_bound' got status invalid (stopping propagation).

- Po přičtení bude hodnota proměnné idx větší než 25. Assert tedy neprošel.
- Kontrola: jaký je zde rozdíl oproti předchozímu případu?

```
idx += 28;
/*@ assert rte: index_bound: 0 <= idx; */
/*@ assert rte: index_bound: idx < 25; */
array[idx] = 15;</pre>
```

hello.c:26:[value] Assertion 'rte,index_bound' got status invalid (stopping propagation).

- ▶ Po přičtení bude hodnota proměnné idx větší než 25. Assert tedy neprošel.
- Kontrola: jaký je zde rozdíl oproti předchozímu případu?
- Následující kód je označen za nedosažitelný. Co to znamená v praxi?

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

7ávěr

Výhody Omezen

Motivace

Value plugin umí dokázat některé vlastnosti funkcí

Motivace

- Value plugin umí dokázat některé vlastnosti funkcí
- Některé ale ne.
- Vyhodnocování složitých smyček je buď nepřesné (odhad invariantu), nebo pomalé (rozbalování).
- WP plugin umožňuje dokazovat vlastnosti programů podobně, jako jsme zvyklí dokazovat vlastnosti algoritmů - deduktivně.
- Silná podpora Design by Contract

Ukázkový kód

Třeba mít nainstalován theorem prover Alt-Ergo

(Vyplatí se prozkoumat dodané Makefily)

Možné načíst pomoci frama-c-gui

- Toto již jsou ručně psané anotace.
- První assert plyne s předchozího IFu.
- WP o stavu programu po skončení cyklu nebo volání funkce nepředpokládá vůbec nic. Proto se nepodaří dokázat druhý assert.

- Toto již jsou ručně psané anotace.
- První assert plyne s předchozího IFu.
- WP o stavu programu po skončení cyklu nebo volání funkce nepředpokládá vůbec nic. Proto se nepodaří dokázat druhý assert.
- Value to zvládne, pokud mu povolíme trávit více času rozbalováním smyček.



S invariantem

- Z invariantu lze snadno dokázat požadované vlastnosti.
- Invariant ovšem platí pouze za podmínky, že nenastane runtime chyba.
- Pokud nelze dokázat RTE assert, invariant nemusí platit.

Podmíněná platnost

```
int array[5];
/*@ loop invariant: \forall integer j;
    0 <= j < i ==> array[j] == 1 - j; */
for (int i = 0; i < 5; i++) {
        i = -1;
/*@ assert rte: index_bound: 0 <= i; */
        array[i] = 1 - i;
}</pre>
```

- ▶ Dojde k runtime chybě (možno zjistit s pomocí Value pluginu)
- WP dokáže totéž, co v předchozím případě.

Terminace

```
int array[5];
/*@ loop invariant: \forall integer j;
    0 <= j < i ==> array[j] == 1 - j; */
for (int i = 0; i < 5; i++) {
        i = 0;
/*@ assert rte: index_bound: 0 <= i; */
        array[i] = 1 - i;
}</pre>
```

- Invariant bezpodmínečně platí (WP ho dokáže)
- Tedy asserty také platí

Terminace

```
int array[5];
/*@ loop invariant: \forall integer j;
    0 <= j < i ==> array[j] == 1 - j; */
for (int i = 0; i < 5; i++) {
        i = 0;
/*@ assert rte: index_bound: 0 <= i; */
        array[i] = 1 - i;
}</pre>
```

- Invariant bezpodmínečně platí (WP ho dokáže)
- Tedy asserty také platí
- Naneštěstí cyklus neterminuje

Správné řešení

```
int array[5];
/*@ loop invariant: 0 <= i <= 5 &&
(    \forall integer j;
0 <= j < i ==> array[j] == 1 - j);
        loop variant 5 - i */
for (int i = 0; i < 5; i++) {
    /*@ assert rte: index_bound: 0 <= i; */
        array[i] = 1 - i;
}</pre>
```

- Omezení na proměnnou i vede k platnosti RTE assertu.
- Hodnota variantu se snižuje v každém kroku cyklu až k nule.
- Program je korektní a terminuje. U rozumných smyček ale není třeba terminaci ověřovat (a psát varianty)

Osnova

Úvod

Jazyk ACSL

Obecné informace Jazykové konstrukce

Použití Frama-C

Příprava zdrojových kódů Value plugin (+ RTE) WP plugin

Závěr

Výhody Omezer

Výhody

- Široká dokumentace
- ► Snadná instalace

► Pouze C (nepodporuje C++)

- ► Pouze C (nepodporuje C++)
- Líbilo by se mi:

- ▶ Pouze C (nepodporuje C++)
- Líbilo by se mi:
 - ► Třídy, používání jmenných prostorů (šetří psaní)

- ▶ Pouze C (nepodporuje C++)
- Líbilo by se mi:
 - Třídy, používání jmenných prostorů (šetří psaní)
- Implementována pouze část ACSL

- ▶ Pouze C (nepodporuje C++)
- Líbilo by se mi:
 - Třídy, používání jmenných prostorů (šetří psaní)
- Implementována pouze část ACSL
- Ošklivé chybové hlášky

Bugy

- Při výpisu chyb nesedí čísla řádků
- Občasný pád grafického rozhraní