# EPFL Formal Verification Course Exam, 23 November 2023

IMPORTANT INFORMATION

**Do not open the exam until we tell you to. The exam is three hours long.**

**Place your CAMIPRO card on your desk.**

**Put all electronic devices in a bag away from bench.**

**Write using permanent, dark pen (no graphite nor heat-disappearing pen).**

**Write answers to different problems on disjoint sheets of paper that we supply.**

**Write your name, SCIPER and question number on the top-right of each sheet you return.**

**Do not write the solutions that you want us to grade on the sheets with exam questions; please take these printed exam sheets with you after the exam.**

**You are only allowed one, A4-sized, two-sided cheat sheet.**

The maximal number of points on the exam is 40. We advise you to first solve questions that you find easier. You can use reasonably high-level mathematical proofs except in part 4.3 where we will need an explicit sequence of proof steps. If you are running out of time on a particular problem, try to convince us that you know the right strategy to solve it.

**Reminder:** If $t \subseteq B \times B$ is a binary relation and $C \subseteq B$, we define

$$t[C] = \{y \mid \exists x \in C.\ (x, y) \in t\}$$

The diagonal relation on $C \subseteq B$ is $\Delta_C = \{(x, x) \mid x \in C\}$. Given $t_1, t_2 \subseteq B \times B$, relation composition $\circ$ is:

$$t_1 \circ t_2 = \{(x, z) \mid \exists y.(x, y) \in t_1 \land (y, z) \in t_2\}$$

Reflexive transitive closure of $t$ is

$$t^* = \bigcup_{i \geq 0} t^i$$

where $t^0 = \Delta_B$, $t^1 = t$, $t^{n+1} = t \circ t^n$. If $M = (S, I, r, A)$ is a transition system we define

$$\bar{r} = \{(s, s') \mid \exists a \in A.(s, a, s') \in r\}$$

and define the reachable states as image of $I$ under the reflexive transitive closure of $\bar{r}$:

$$Reach(M) = \bar{r}^*[I]$$

**Do not open the exam until we tell you to.**

# 1 Transition Systems and Invariants (8pt)

Let $M = (S, I, r, A)$ be a transition system with $I \subseteq S$ and $I \neq \emptyset$ a non-empty initial set of initial states, $r \subseteq S \times A \times S$ the transition relation and $A$ the (non empty) input alphabet.

**Part I.** For each of the following, prove or give a counter-example.

## 1.1 Diagonal

If $\bar{r}$ is the diagonal relation, then every subset of $S$ is an inductive invariant.

## 1.2 Lurking Havoc

If there exists an $s_1 \in S$ such that for all $s_2 \in S$, $(s_1, s_2) \in \bar{r}$, then the smallest inductive invariant is $S$.

## 1.3 Rock Star

For any $s_2$: if for all $s_1 \in S$ we have $(s_1, s_2) \in \bar{r}$, then $s_2 \in \mathit{Reach}(M)$. Stated as a formula:

$$\forall s_2. ((\forall s_1 \in S.(s_1, s_2) \in \bar{r}) \longrightarrow s_2 \in \mathit{Reach}(M))$$

## 1.4 Transitions over a Lattice

If $S$ is a complete lattice with the order relation $\leq$ and with the greatest lower bound of a set denoted by $\bigsqcap$, then if for all $s_1, s_2$, $(s_1, s_2) \in \bar{r} \to s_1 \leq s_2$, then for $u = \bigsqcap I$, the set $I_u = \{s \in S \mid u \leq s\}$ is an inductive invariant.

## 1.5 No Junk

If, for all $s \in S \setminus I$, $\bar{r}[\{s\}] \subseteq \mathit{Reach}(M)$, then every invariant is an inductive invariant.

****

**Part II.** In the following parts, for any set of initial states $I' \subseteq S$ and relation $r' \subseteq S \times A \times S$, let $M(I', r')$ denote the transition system $(S, I', r', A)$ and let $\mathit{Reach}(I', r')$ denote $\mathit{Reach}(M(I', r'))$. For each of the following, prove or give a counter-example.

## 1.6 One Reach, Two Reaches

For any $r_1$, $r_2$ and $I$, $\mathit{Reach}(\mathit{Reach}(I, r_1), r_2) \subseteq \mathit{Reach}(I, r_1 \cup r_2)$.

## 1.7 Both Reach

For any $r_1$, $r_2$ and $I$, $\mathit{Reach}(I, r_1 \cup r_2) \subseteq \mathit{Reach}(\mathit{Reach}(I, r_1), r_2)$.

### 1.8 Invariant Part by Part

For any $r_1$, $r_2$ and $I$, if $i_1$ is an invariant of $M(I, r_1)$ and $i_2$ is an invariant of $M(I, r_2)$ then $i_1 \cup i_2$ is an invariant of $M(I, r_1 \cup r_2)$.

## 2 Quantifier Elimination (5pt)

These problems can be solved using the technique of quantifier elimination for linear arithmetic that we studied in the lectures. You can skip steps when transforming quantifier-free sub-formulas into equivalent ones. You can use your own strategy for applying quantifier elimination steps.

### 2.1 An Integer Formula

Consider the formula $F(x, z)$ where the variables range over the mathematical integers $\mathbb{Z}$:

$$\forall y.((x < y \land y < z) \longrightarrow (3 \mid (y+1) \lor 3 \mid (y+2)))$$

($3 \mid (y+1)$ stands for "3 divides $(y+1)$".)

Find a quantifier-free formula equivalent to $F(x, z)$.

### 2.2 A Rational Arithmetic Formula

Consider the following formula $G(x, z)$ where the variables range over rational numbers $\mathbb{Q}$:

$$\forall y.((x < y \land y < z) \longrightarrow \forall u.(x \neq u + u + u))$$

Find a quantifier-free formula equivalent to $G(x, z)$.

More questions on next pages!

# 3 Programs and Formulas (9pt)

Consider a program with two variables ranging over unbounded integers $\mathbb{Z}$, so let $S = \mathbb{Z}^2$ be the set of possible states of the program at a program point.

## 3.1 Program to Formula

Convert the following program into a formula that expresses the program's meaning:

```
x= x + y
if x > y then
  x= x - y
else
  y= y - x
```

Your formula should have:

1. $x, y, x', y'$ as the only *free* variables
2. "$\exists$" as the only quantifier (no universal quantifiers); it may also be quantifier-free
3. no relations or functions other than those of integer linear arithmetic.

## 3.2 About That Program

Let $r \subseteq S \times S$ be the semantics of the program in the previous part. Consider the initial set of states $p$:

$$p = \{(x, y) \in S \mid 0 \leq x + y\}$$

Find the formula that describes the strongest postcondition of $p$ with respect to $r$. Specifically, compute a quantifier-free formula $Q$ containing as the variables only $x$ and $y$, such that the set

$$q = \{(x, y) \mid Q\}$$

is the relation image of set $p$ under relation $r$, that is, $q = r[p]$ holds.

You do not need to show detailed steps, but be careful to give a formula for the exact relation image (strongest postcondition).

## 3.3 Quadratic Mess

This part is solvable independently of the previous two. Consider the initial set of states

$$p = \{(x, y) \in S \mid x^2 \leq y\}$$

Let $r \subseteq S \times S$ now be the meaning of the following program with two assignments executed one after another:

```
y= y − 1
x= x + y
```

Compute a quantifier-free formula $Q$ containing as variables only $x$ and $y$ that characterizes the strongest postcondition of $p$ with respect to $r$, that is, a formula $Q$ such that the set

$$q = \{(x, y) \mid Q\}$$

is the relation image of set $p$ under relation $r$, that is, $q = r[p]$ holds.

# 4 First Order Logic Resolution (7pt)

Consider the following first order logic signature $L = \{E, a, f, g\}$, consisting of a binary relation symbol $E$, a constant $a$, a function $f$ taking 1 argument, and a function $g$ taking 1 argument. Recall that, in FOL, we distinguish terms (denoting values in the domain) from formulas (denoting truth values).

## 4.1 Defining Terms

Let $V$ denote a countable set of variables used to build terms and formulas. Let $T$ denote the set of all terms with variables $V$ in the signature $L$. Let $\mathsf{GT}$ denote the set of ground terms corresponding in the signature $L$ (domain of the Herbrand universe).

a) Does $g(f(a))$ belong to $\mathsf{GT}$?

b) Does $g(f(x))$ belong to $\mathsf{GT}$?

c) Does $E(f(a), a)$ belong to $\mathsf{GT}$?

d) Does $E(f(x), a)$ belong to $T$?

e) Give a definition for a function $H : 2^T \to 2^T$ such that these two conditions hold:

$$T = \bigcup_{i \geq 0} H^i(V)$$

$$\mathsf{GT} = \bigcup_{i \geq 0} H^i(\emptyset)$$

Do not use $\mathsf{GT}$ in the definition of $H$. (This way, we can use $H$ to define $\mathsf{GT}$.)

($H^i(x)$ denotes the iterated application of $H$, i.e., $H^0(x) = x$, $H^{i+1}(x) = H(H^i(x))$.)

More questions on next pages!

### 4.2 Axioms and Their Normal Form

Consider the following formula $A$:

$$\begin{array}{ll}\forall x, y, z. & (E(x,y) \land E(y,z) \to E(x,z)) \land \\ & (E(x,y) \to (E(f(x), f(y)) \land E(g(x), g(y)))) \land \\ & E(f(g(x)), g(f(x)))\end{array}$$

Show the result of transforming the above formula into an *equivalent* finite set of first-order clauses.

### 4.3 Applying Resolution

Use the clauses obtained in the previous part to show that $E(f(f(g(a))), g(f(f(a))))$ is a consequence of formula $A$.

Use a refutation proof with the rule of FOL resolution with instantiation. Write your proof as a numbered sequence proving the empty clause $\emptyset$. For each step indicate if it is an assumption or write "from $n_1$, $n_2$" where $n_1$ and $n_2$ are previous steps from which it follows.

You may abbreviate the terms using prefix notation, writing e.g. $E(f(f(g(a))), g(f(f(a))))$ as $E(\mathit{ffga}, \mathit{gffa})$.

## 5 The Age of AI - Abstract Interpretation (11pt)

In this question we are designing an abstract domain that improves on intervals by tracking divisibility as well. Even in case your understanding of abstract interpretation is limited, your intuition and understanding of Hoare triples and strongest postconditions may allow you to solve some of these problems.

For simplicity, consider programs with a single variable. A set of states is a subset of the set of integers, so the concrete domain is the set of all subsets, $C = 2^{\mathbb{Z}}$. The ordering on concrete elements is $\subseteq$, which gives a (complete) lattice with the least upper bound $\cup$ and the greatest lower bound $\cap$.

The abstract domain elements are four-tuples $(a, b, c, d)$ where $a, b, c, d$ can be integers and where $a$ can also be $-\infty$ and $b$ can also be $+\infty$. We assume that $-\infty \leq x$ and $x \leq +\infty$ for every $x \in \mathbb{Z}$. Hence,

$$A = \{(a, b, c, d) \mid a \in \{-\infty\} \cup \mathbb{Z}, b \in \mathbb{Z} \cup \{+\infty\}, c, d \in \mathbb{Z}\}$$

Define

$$\gamma(a, b, c, d) = \{x \mid a \leq x \leq b \land \exists k \in \mathbb{Z}.\ x = kd + c\}$$

### 5.1 Special Case

Give a simple definition of the set $\gamma(a, b, 0, 1)$ and for the set $\gamma(a, a, 0, 1)$.

## 5.2 Abstract Strongest Postcondition for Assignment

Consider the following assignment statement $c1$:

```
x = x - 3
```

Let $r_1 \subseteq \mathbb{Z}^2$ be the meaning of that statement. Write down an expression defining $r_1$.

Then, give a definition of a function $F_1^\# : A \to A$ such that, for all $x \in A$,

$$r_1[\gamma(x)] \subseteq \gamma(F_1^\#(x)) \tag{1}$$

Try to define $F_1^\#$ such that $\gamma(F^\#(x))$ is as small set as possible while satisfying the above condition.

Illustrate your definition by showing and simplifying the mathematical expression for $F^\#((a, a, 0, 1))$ where $a \in \mathbb{Z}$.

## 5.3 Abstract Strongest Postcondition for Tests

Analogously to the previous part, consider the command $c2$:

```
assume(x > 3)
```

whose meaning is relation $r_2$. Give functions $F_2^\# : A \to A$ that satisfies the condition analogous to (1):

$$r_2[\gamma(x)] \subseteq \gamma(F_2^\#(x))$$

Also give $F_3^\# : A \to A$ that corresponds to the command:

```
assume(x <= 3)
```

## 5.4 Joining

Propose a definition of $J : A \times A \to A$ such that for all $x, y$:

- $J(x, y) = J(y, x)$
- $\gamma(x) \subseteq \gamma(J(x, y))$

and such that $\gamma(J(x, y))$ is as small as you can make it while satisfying the above two conditions.

More questions on next pages!

## 5.5 Loop and Its Control-Flow Graph

You may be able to solve this part independently of the other parts.

Consider the following small program.

```
// 1
x = 20
while // 2
      x > 3 do
      // 3
     x = x − 3
// 4
```

Draw a control flow diagram with these 4 program points, with edges labeled by assignments and tests ("assume" statements).

Let $I_1$ be the mathematical formula "true".

Find formulas $I_2, I_3, I_4$ expressible in the language of integer linear arithmetic with the divisibility operator such that $I_1, I_2, I_3, I_4$ result in valid Hoare triples according to the control flow graph, and such that they are as strong as you can make them (making all of them "true" is not a good solution).

For example, the following should be valid Hoare triples (among others):

- $I_1$ {`x = 20`} $I_2$
- $I_2$ {`assume(x > 3)`} $I_3$
- $I_3$ {`x = x - 3`} $I_2$

## 5.6 Injectivity

Give an example showing that $\gamma$ is not injective.

Define a subset $A_N \subseteq A$ such that $\gamma$ restricted to $A_N$ is injective and $\gamma[A_N] = \gamma[A]$.

We write $\gamma_N$ for the restriction of $\gamma$ to $A_N$: $\gamma_N(x) = \gamma(x)$ for all $x \in A_N$.

## 5.7 Ordering on $A$

Define a binary partial order relation $\leq$ on $A_N$ such that $\gamma_N$ is an injective monotonic function from $A_N$ to $C$.

## 5.8 Galois Insertion

Can you define $\alpha : C \to A_N$ such that $(\alpha, \gamma_N)$ form a Galois insertion between $C$ and $A_N$? (Reminder: a Galois insertion is a Galois connection where $\gamma_N$ is injective.)

**End of the Exam Sheet**