Software Verification

Friday, January 8th 2016, 3 hours

This assignment contains three independent parts: the first part deals with bounded model-checking, the second part is about Galois connections and range analysis, and the last part addresses Craig interpolation.

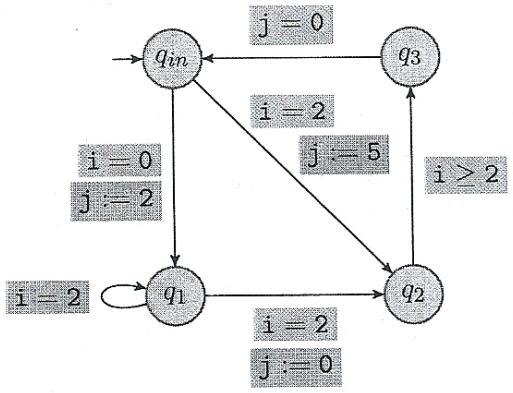
All documents are authorized during the examination.

1 Bounded Model-Checking (8pts)

This section will browse a few concepts that have been seen in the first part of the course (Bounded Model-Checking).

1.1 Accessibility in a finite states program

Question 1 *Given the following program P, list all the possible configurations that can be reached by P starting with* $\texttt{i} = 0$, $\texttt{j} = 0$ *in* q_{in}.



1.2 Binary Decision Diagrams

Question 2 *Given* $A = (x \wedge y) \vee (\neg x \wedge z) \vee (\neg y \wedge z)$ *and* $B = (x \wedge y \wedge t) \vee (\neg x \wedge \neg z \wedge t)$, *two propositional logic formula, build the resulting binary decision diagrams for A, B and* $(A \vee B) \wedge (\neg A \wedge B)$ *(order on variables is* $x > y > z > t$*).*

1.3 Merry Christmas 2015 from the GCHQ!

This year, the greeting Christmas card from the GCHQ (Fig. 1) was a Nonogram displaying a QR-code once solved. The point of this exercise is to use a SAT-solver to find the solution.

Note that we are interested in finding a way to solve the problem and we do not ask the solution of Fig. 1, so **do not try to solve it during the exam!**

The rules for a Nonogram is that you are given for each row and column the number of consecutive cells to shade. For example, if you are given: 7, 5, 4 for a row, it means

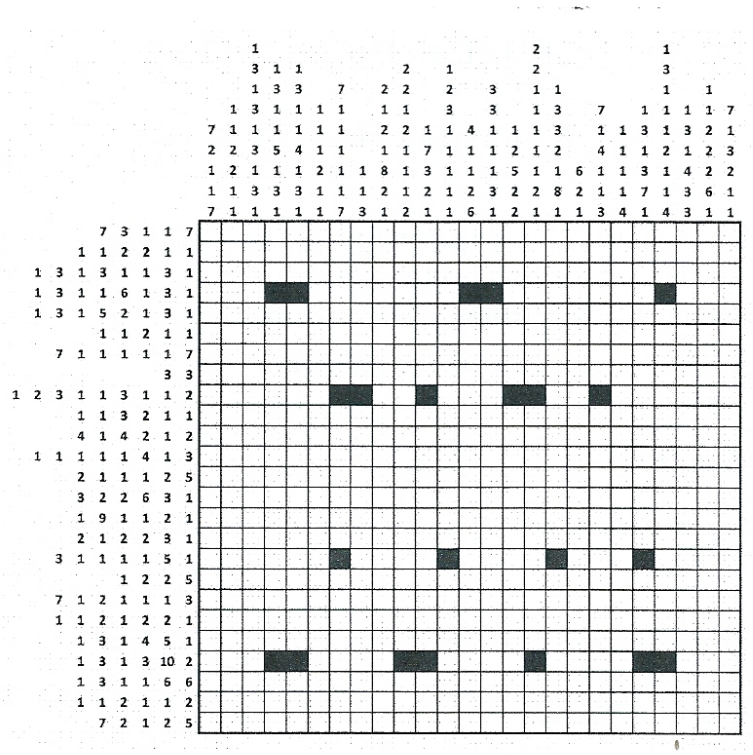


Figure 1: Christmas 2016 greeting card from GCHQ

that the row hold a sequence of 7 consecutive shaded cells, then one or more white cells, then 5 consecutive shaded cells, and so on. The grids are designed to have only one unique solution. But, in order to force one unique solution or to give a few clues to the player, a few pre-shaded cells may be added to the grid.

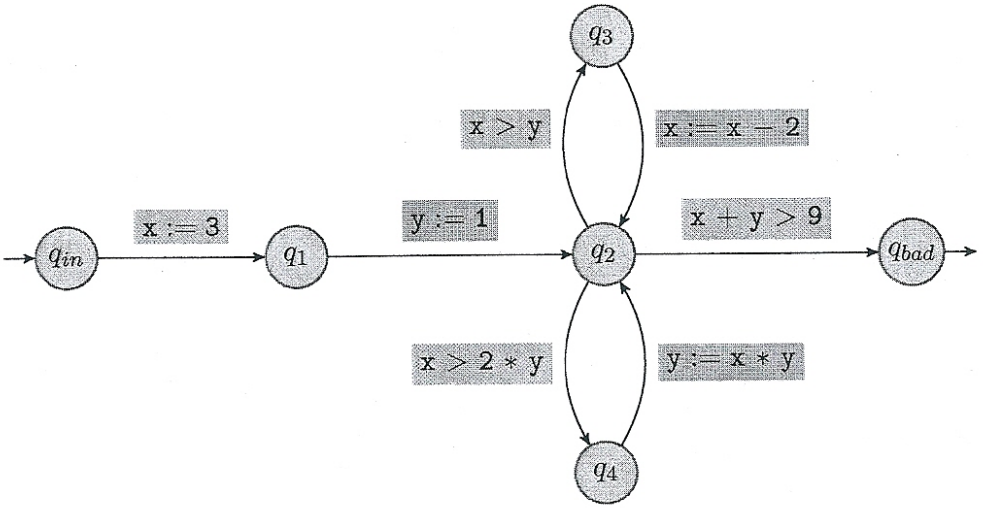
Variables To transform this puzzle into a propositional formula, we consider the boolean variables $cell_{i,j}$ (where i is the row number and j the column number). The cell at the position (i, j) is shaded if and only if $cell_{i,j}$ is $True$. In our example Fig. 1, the grid is of size 25×25 which makes 625 variables. Then, for a given row i, we call $rowseq_{i,k}$ the size of the k^{th} sequence of this row. And, for a given column j, we call $colseq_{j,k}$ the size of the k^{th} sequence of this column.

Hint There are several ways to solve this problem with propositional logic, but one good idea would be to introduce extra variables to represent all the possible positions of each sequence. The variable $rowseqpos_{i,k,p}$ is $True$ if and only if the k^{th} sequence of the row i is starting at column p.

Question 3 *Give the structure of a propositional logic formula that can be used to solve this problem with a SAT-solver. Explain the extra boolean variables that you may introduce and the way you encode the constraints of the problem.*

2 Range Analysis (4pts)

We perform range analysis on the control-flow automaton depicted below, with variables $\mathtt{X} = \{\mathtt{x}, \mathtt{y}\}$, both ranging over integers. The initial location is q_{in} and bad location is q_{bad}.



Like in the course, an analysis will be called *successful* when the abstract value obtained for q_{bad} is $\bot$. Round-robin iteration shall use the following order on locations: $q_{in}, q_1, q_2, q_3, q_4, q_{bad}$.

Question 4 *Apply the round-robin algorithm with widening. Do* not *use narrowing. Is the analysis successful?*

Question 5 *Starting from the result of the previous question, perform a decreasing iteration with narrowing. Is the analysis successful?*

3 Craig Interpolation (8pts)

In this section we consider 2-SAT formulas. A 2-SAT formula is a conjunction of clauses that contains only two literals.

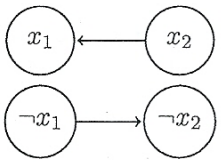
Question 6 *Provide a resolution tree proving that the 2-SAT formula* $(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee \neg x_2) \wedge (x_2 \vee x_3) \wedge (x_2 \vee \neg x_3)$ *is unsat.*

Question 7 *From the previous resolution tree, provide an interpolant for the decomposition of the formula into a left part* $(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee \neg x_2)$ *and a right part* $(x_2 \vee x_3) \wedge (x_2 \vee \neg x_3)$.

We associate to a 2-SAT formula ϕ the directed graph $G_\phi = (V_\phi, \to_\phi)$ defined as follows:

- V_ϕ is the set of literals x or $\neg x$ where x ranges over the variables occuring in ϕ, and
- $\to_\phi$ is a binary relation over V_ϕ defined by $l \to_\phi l'$ if $\neg l \vee l'$ is a clause of ϕ (with the convention $\neg\neg x = x$).

Example 1 *The directed graph associated to the formula* $(x_1 \vee \neg x_2)$ *is the following one:*



Question 8 *Draw the directed graph associated to* $(x_1 \vee \neg x_2) \wedge (\neg x_1 \vee \neg x_2) \wedge (x_2 \vee x_3) \wedge (x_2 \vee \neg x_3)$.

We introduce the binary relation $\xrightarrow{*}_\phi$ over the literals in V_ϕ defined by $l \xrightarrow{*}_\phi l'$ if there exists a directed path in G_ϕ from l to l'.

Question 9 *Prove that if* $l \xrightarrow{*}_\phi l'$ *then every valuation satisfying* ϕ *satisfies* $l \Rightarrow l'$. *Tips : Induction on the length of paths.*

Question 10 *Prove that if* $l \xrightarrow{*}_\phi l'$ *then* $\neg l' \xrightarrow{*}_\phi \neg l$. *Tips : Induction on the length of paths.*

Question 11 *Prove that if* ϕ *is a 2-SAT formula such that for some variable* x, *we have* $x \xrightarrow{*}_\phi \neg x$ *and* $\neg x \xrightarrow{*}_\phi x$, *then* ϕ *is unsat. Tips : Proof by contradiction.*

In fact, the converse is true, i.e. if for every variable x we do not have $x \xrightarrow{*}_\phi \neg x$ and $\neg x \xrightarrow{*}_\phi x$ then ϕ is sat.

Question 12 *Deduce a simple polynomial time algorithm for deciding the satisfiability of a 2-SAT formula* ϕ *(just provide the idea of the algorithm).*

Now, let ϕ_L and ϕ_R be two 2-SAT formulas, and let us introduce $\phi = \phi_L \wedge \phi_R$.

Question 13 *Assume that there exists a variable* x *such that* $x \xrightarrow{*}_\phi \neg x$ *and* $\neg x \xrightarrow{*}_\phi x$. *Prove that:*

- *Either* $x \xrightarrow{*}_{\phi_L} \neg x$ *and* $\neg x \xrightarrow{*}_{\phi_L} x$,
- *Or* $x \xrightarrow{*}_{\phi_R} \neg x$ *and* $\neg x \xrightarrow{*}_{\phi_R} x$,
- *Or there exists a global variable* $y \in V_{\phi_L} \cap V_{\phi_R}$ *such that* $y \xrightarrow{*}_\phi \neg y$ *and* $\neg y \xrightarrow{*}_\phi y$.

Tips : If the paths $x \xrightarrow{*}_\phi \neg x$ *and* $\neg x \xrightarrow{*}_\phi x$ *contains both an edge* $\rightarrow_{\phi_L}$ *and* $\rightarrow_{\phi_R}$, *then we can find a global variable* $y \in V_{\phi_L} \cap V_{\phi_R}$ *that occurs along one of the two paths* $x \xrightarrow{*}_\phi \neg x$ *or* $\neg x \xrightarrow{*}_\phi x$. *Consider the two cases separately.*

From the previous question, it is possible to deduce a simple algorithm for computing Craig interpolants for (ϕ_L, ϕ_R) when $\phi_L \wedge \phi_R$ is unsat.