INFOMLSAI Logics for Safe AI
Coursework 2

Coursework released:	3 December 2024, on Blackboard
Coursework due:	23:59 17 December 2024, on Blackboard
Submission format:	a folder containing a pdf and an `.ispl` file, one per group

Please do the coursework in groups of 3-4 people. Submit a single folder on Blackboard for your group, indicating in the pdf file and in comments in the ispl file who the members of the group are.

W2-0 To prepare for the model-checking part of the coursework, read the guide to installing MCMAS on Blackboard, and download and install MCMAS. Download the file `one-robot+carriage.ispl` from Blackboard. Check that everything works correctly by running MCMAS with `one-robot+carriage.ispl` as input.

The topic of the coursework is inspired by the autonomous underwater vehicle verification project described in [1] but of course we keep it a lot simpler.

Consider a submarine whose state is described by boolean variables surface (meaning, the submarine is on the surface), open (the hatch is open), sunk (self explanatory). The submarine can perform actions $\{up, down, open, close\}$. There are 4 possible states. In the initial state s_1, the submarine is on the surface, not sunk, and the hatch is closed. Action $open$ moves it into state s_2 where it is still not sunk and on the surface and the hatch is open. Action $down$ from s_1 results in s_3 where the submarine is not on the surface, not sunk, and the hatch is closed. Action $down$ from s_2 results in s_4 where the submarine is not on the surface, the hatch open, and the submarine is sunk.

If the submarine is on the surface and performs up, then it remains on the surface; if it is not on the surface and performs up, then next it will be on the surface. Similarly, $down$ does nothing if the submarine is not on the surface, otherwise surface becomes false. The action $open$ makes open true; the action $close$ makes open false. If $open$ is performed when the submarine is not on the surface, sunk becomes true. If open is true and the submarine performs $down$, then sunk becomes true. Once sunk is true, the only possible action is $down$ and it does not change the state of the submarine.

CW2-1 Define the system above formally (specify $St, \longrightarrow, \mathcal{V}$).

(1 point)

CW2-2 Consider the state s_1 in your state transition system where the submarine is on the surface, and propositions open and sunk are false. Express in CTL: on all paths, always, the submarine is not sunk. Is this property true in s_1? Explain your answer with the reference to CTL truth definitions.

(1 point)

CW2-3 Express in CTL: there exists a path where in some future state the submarine is not on the surface and not sunk, and until that state, it holds that the submarine was also not sunk. Is this property true in s_1? Explain your answer with the reference to CTL truth definitions.

(1 point)

CW2-4 The semantics for CTL given in the lectures assumes all paths are infinite. This is not completely realistic. For example, in the submarine scenario, it would be more natural to say that there are no actions available after the submarine is sunk, and the state where it is sunk has no successors, i.e., it is the last state on the path. There is a version of CTL called CTL on finite traces, CTL_f, where paths are finite. In the truth definition for CTL_f, the clause for EX is:

$M, q \models EX\varphi$ if, and only if, there is a path $q, q_1, \ldots$ from q, and $M, q_1 \models \varphi$.

Is there any difference between the set of formulas that are true in all CTL models and the set of formulas true in all CTL_f models? Explain why. If there is a difference, give a formula which is true in all CTL models but not in CTL_f models (or vice versa).

(1 point)

CW2-5 Create a description in ISPL of the submarine agent. The initial state is s_1 (on the surface, hatch closed, not sunk).

(2 points)

CW2-6 The following statements should be true in the initial state. Translate them to MCMAS notation and add them to your file.

- On all paths there is a future state where surface is true and open and sunk are false (this just describes the initial state, since it is included in all paths starting from it)
- On all paths in all states, if surface is true then it is possible to make it false in the future
- On all paths in all states, if open is false then it is possible to make it true in the future
- There is a path where from some future state, there is a path where in every state the submarine is not on the surface but is also not sunk.

Extract a witness for the last formula. (1 point)

CW2-7 The following statements should be false in the initial state. Translate them to MCMAS notation and add them to your file.

- On all paths globally the hatch is never open.
- On all paths at some point in the future the submarine is sunk.
- There is a path where at some point the submarine is sunk and from there there is a path where at some point in the future it is not sunk.
- On all paths in the next state it holds that on all paths in the next state the submarine is not sunk.

Extract a counterexample for the last formula. (1 point)

CW2-8 The CTL truth definition for $E\varphi\, U\, \psi$ does not require φ to hold in the state satisfying ψ, but only in preceding states. Consider an alternative definition $M, q \models E\varphi U^{+}\psi$ iff there exists a path λ from q such that for some $i \geq 0$, $M, \lambda[i] \models \psi$ and for all j, $0 \leq j \leq i$, $M, \lambda[j] \models \varphi$.

Write the case for the model checking algorithm for $E\varphi U^{+}\psi$ modality.

(1 point)

References

[1] Jonathan Ezekiel, Alessio Lomuscio, Levente Molnar, and Sandor M. Veres. Verifying fault tolerance and self-diagnosability of an autonomous underwater vehicle. In Toby Walsh, editor, *IJCAI 2011, Proceedings of the 22nd International Joint Conference on Artificial Intelligence*, pages 1659–1664. IJCAI/AAAI, 2011.