Time-Staging Enhancement of Hybrid System Falsification

Gidon Ernst[1], Ichiro Hasuo[1], Sean Sedwards[2], Zhenya Zhang[1]
[1] *National Institute of Informatics, Tokyo, Japan*
[2] *University of Waterloo, Waterloo, Canada*

Problem

- Falsification problem is defined as follows:
  - **Given:** a *model* $\mathcal{M}$ (that takes an input signal $\mathbf{u}$ and yields an output signal $\mathcal{M}(\mathbf{u})$), and a *specification* φ (a temporal formula)
  - **Answer:** *error input*, that is, an input signal $\mathbf{u}$ such that the corresponding output $\mathcal{M}(\mathbf{u})$ violates φ
- Challenges:
  - Black/Grey box model, e.g., model in Simulink, etc.
  - Continuous (infinite) input space

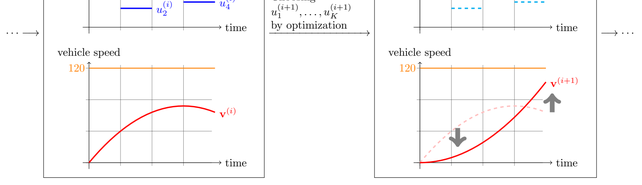
$$\xrightarrow{\mathbf{u}} \boxed{\mathcal{M}} \xrightarrow{\mathcal{M}(\mathbf{u}) \quad \not\models \varphi\,?}$$

Related work

- Robust semantics of temporal formulas
  - Traditional: Boolean satisfaction relation $\mathbf{v} \models \varphi$
  - Now: quantity $[\![\mathbf{v}, \varphi]\!] \in \mathbb{R} \cup \{\infty, -\infty\}$ $[\![\mathbf{v}_2, \varphi]\!] = -10$
- Optimization-based falsification:
  - Objective function: $[\![\mathbf{v}, \varphi]\!]$
  - Solvers: Nelder-Mead, CMA-ES, Simulated Annealing, etc. $[\![\mathcal{M}(\mathbf{u}^{(i+1)}), \varphi]\!] > [\![\mathcal{M}(\mathbf{u}^{(i)}), \varphi]\!]$

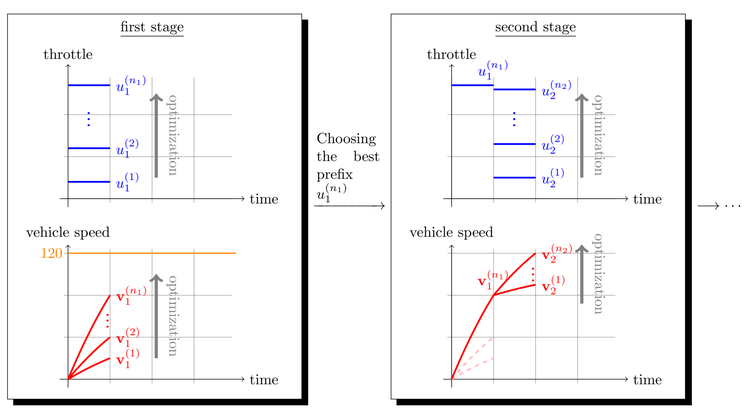
Motivating example

- $\square$ (*speed* < 120)



- In the i-th sampling one tries an input signal $\mathbf{u}^{(i)}$. The corresponding output signal $\mathbf{v}^{(i)} = \mathcal{M}(\mathbf{u}^{(i)})$ is shown below it.
- Optimization algorithm decides a new input signal $\mathbf{u}^{(i+1)} = (u_1^{(i+1)}, \ldots, u_K^{(i+1)})$. and $\mathbf{u}^{(i+1)}$ makes the robustness smaller (i.e. the peak vehicle speed higher).

Time-staged strategy



- In the first stage (left), we run a falsification algorithm and try to find an initial input segment that achieves low robustness (i.e. high peak speed).
- This process will gradually improve candidates for the initial input segment, in the way the arrows ↑ on the left in the figure

Time-staged falsification algorithm

Require: a falsification solver Falsify, a system model $\mathcal{M}$, an **STL** formula φ, $T \in \mathbb{R}_{>0}$ and $K \in \mathbb{N}$

1: $\mathbf{u} \leftarrow ()$ ▷ the input prefix obtained so far. We start with the empty signal ()
2: **for** $j \in \{1, \ldots, K\}$ **do**
3: $\quad \mathbf{u}' \leftarrow \mathsf{Falsify}(\mathcal{M}_{\mathbf{u}}, \partial_{\mathcal{M}(\mathbf{u})}\rho_{\varphi}, \frac{T}{K})$ ▷ synthesizing the j-th input segment
4: $\quad \mathbf{u} \leftarrow \mathbf{u} \cdot \mathbf{u}'$ ▷ concatenate $\mathbf{u}'$, after which the length of $\mathbf{u}$ is $\frac{jT}{K}$
5: **return u** ▷ a time-staged falsification trial is successful if $[\![\mathcal{M}(\mathbf{u}), \varphi]\!] < 0$

Benchmark 1: Automatic transmission

	S1		S2		S3 easy		S3 hard		S4 easy		S4 hard	
Algorithm	time	#/20	time	#/20	time	#/20	time	#/20	time	#/20	time	#/20
CMA-ES	27s	**20**	**5s**	**20**	39s	14	57s	0	32s	16	59s	0
+TS	52s	15	15s	16	**9s**	19	23s	**11**	15s	14	24s	3
+A-TS	41s	18	15s	17	**9s**	16	21s	**10**	26s	14	**20s**	**5**
SA	50s	5	43s	7	37s	9	55s	0	35s	6	47s	**5**
+TS	37s	**20**	33s	**16**	**11s**	**19**	33s	**8**	21s	14	51s	0
+A-TS	34s	**20**	18s	**17**	**9s**	**18**	26s	**4**	**16s**	**18**	30s	2
GNM	**6s**	**20***	61s	0*	56s	0*	55s	0*	43s	0*	53s	0*
+TS	42s	20*	15s	**20***	**13s**	**20***	25s	**20***	**11s**	**20***	52s	0*
+A-TS	20s	20*	16s	**20***	**10s**	**20***	26s	**20***	13s	20*	43s	0*

S1 $\square_{[0,30]}$ $(speed < 120)$
S2 $\square_{[0,30]}$ $(gear = 3 \rightarrow speed \geq 30)$
S3 $\Diamond_{[10,30]}$ $(speed \leq v_{\min} \vee speed \geq v_{\max})$, easy: $v_{\min} : 50$, $v_{\max} : 60$; hard: $v_{\min} : 53$, $v_{\max} : 57$.
S4 $\square_{[0,10]}(v_{\min} < speed) \vee \Diamond_{[0,30]}(rpm > \omega_{\max})$ easy: $v_{\min} : 80, \omega_{\max} : 4500$; hard: $v_{\min} : 50, \omega_{\max} : 2520$.

Benchmark 2: Abstract fuel control

	$\neg(\Diamond_{[0,6]}\square_{[0,3]}(AF - AF_{\text{ref}} > 0.07 * 14.7))$		$\neg(\Diamond_{[6,26]}\square_{[0,4]}(AF - AF_{\text{ref}} > 0.01 * 14.7))$	
Algorithm	time	#/20	time	#/20
CMA-ES	49s	0	82s	1
+TS	30s	0	42s	1
+A TS	26s	0	41s	0
SA	51s	0	76s	2
+TS	47s	1	54s	7
+A TS	34s	0	42s	5
GNM	50s	0*	86s	0*
+TS	30s	20*	20s	20*
+A TS	37s	0*	19s	20*

Conclusion & Future work

We have introduced and evaluated the idea of time staging to enhance falsification for hybrid systems. The proposed method emphasizes exploitation over exploration as part of stochastic optimization. As there is no single algorithm that fits every problem (as a consequence of having no free lunch), having a variety of methods at disposal permits the user of a system to choose the one suitable for the problem at hand. We have shown that the proposed approach is a good fit for problems that suitable exhibit time-causal structures, where it significantly outperforms non-staged algorithms.

Three directions for future work:

- Instead of just picking the best trajectory for each stage, it might be beneficial to retain a few, potentially diverse ones.
- Discover time stages adaptively.
- Explore variations of robust semantics to mitigate discrete propositions.

Acknowledgements

This work is supported by ERATO HASUO Metamathematics for Systems Design Project (No. JPMJER1603), Japan Science and Technology Agency.