Fine-Tuning Language Models Using Formal Methods Feedback: A Use Case in Autonomous Systems

Yunhao Yang*, Neel P Bhatt*, Tyler Ingebrand*, William Ward, Steven Carr, Zhangyang Wang, Ufuk Topcu (* Equal Contribution)
The University of Texas at Austin, USA

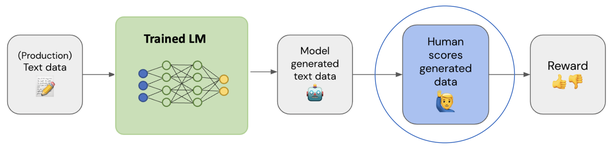
The Central Question

How can we integrate multimodal pretrained models into the algorithms for **verifiable** sequential decision-making?

Problems of Reinforcement Learning from Human Feedbacks…

1) Labor-intensive due to excessive human-annotated data.

Labor-Intensive



2) Human feedbacks is often inconsistent due to their preferences and knowledge.

Subjective/Inconsistent Feedback

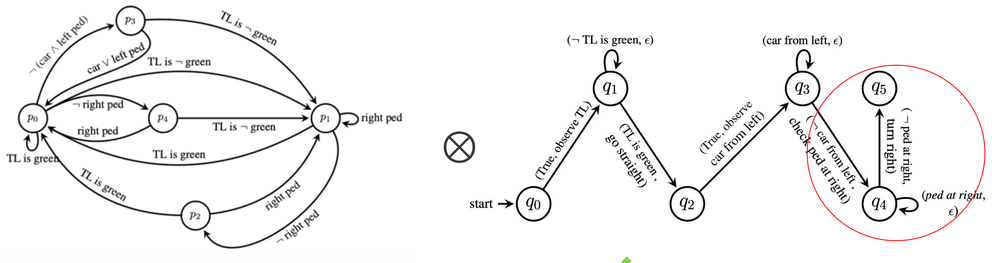
Motivation

How can we fine-tune a large language model for domain specific tasks, e.g., autonomous driving, **without the need for human experts**?
How can we automatically generate **unlimited and consistent** training data when fine-tuning the language model?
How can we check whether the language model's outputs satisfy the autonomous system's requirements.

Contributions

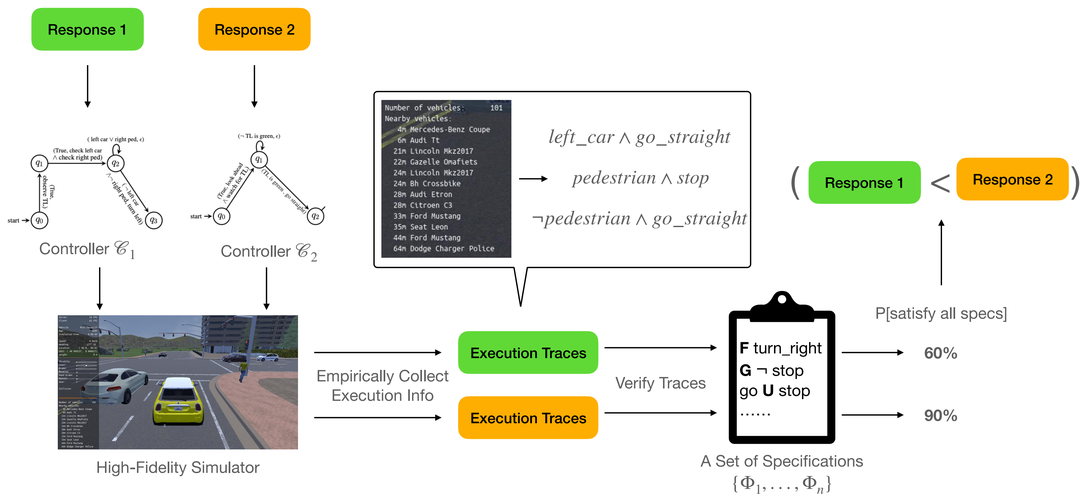
1. Use **formal methods** to **provide feedbacks** to the language model's outputs, eliminate the need for human labeling.
2. Generate and verify task controllers to **ensure consistencies** with **the autonomous system's requirements**.
3. Develop a method that provides automated feedbacks either through formal verification or through empirical data obtained from simulations.

Formal Verification

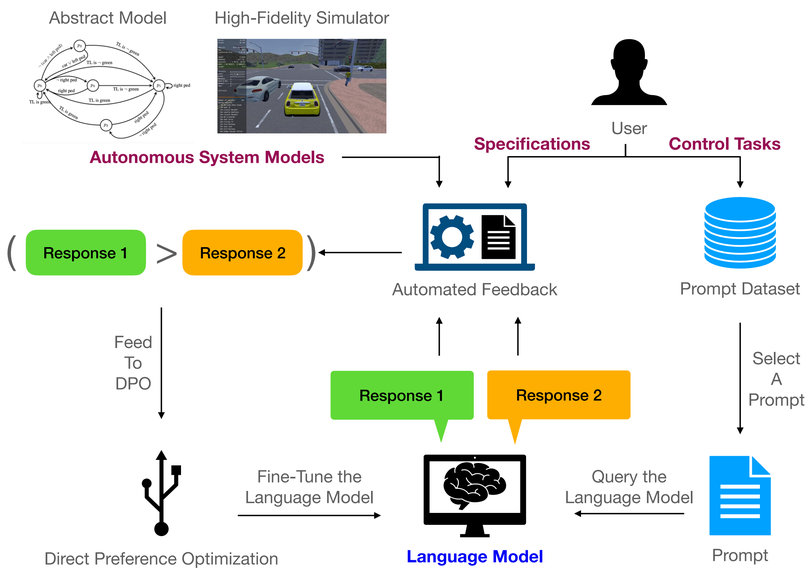


$\square(\neg \text{green traffic light} \rightarrow \neg \text{go straight})$, ✓
$\square(\text{stop sign} \rightarrow \Diamond\, \text{stop})$, ✓
$\square\, \neg \text{turn right} \vee \neg(\text{car from left} \vee \text{pedestrian at right})$, ✗

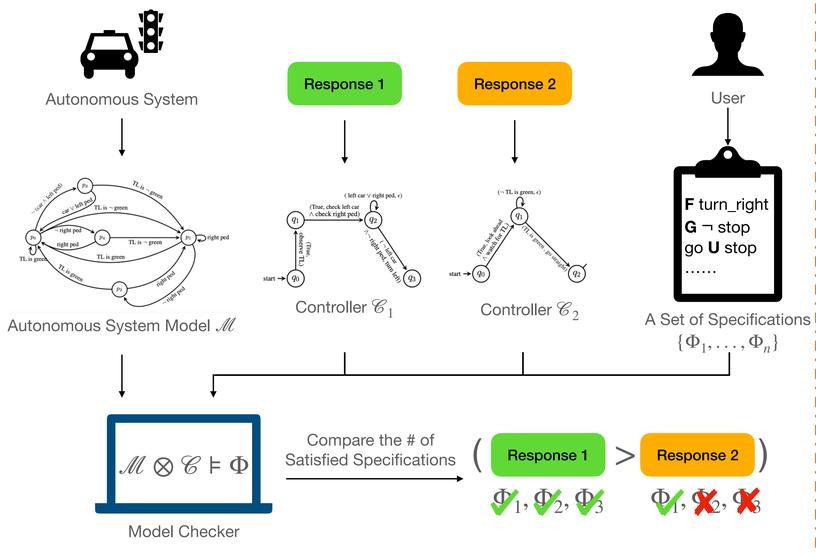
Empirical Evaluation via Simulation



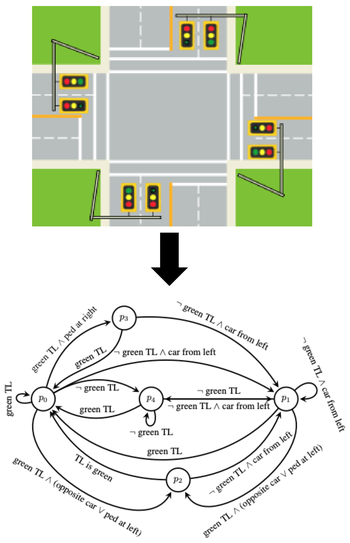
The Overall Pipeline



How do we collect formal methods feedback?



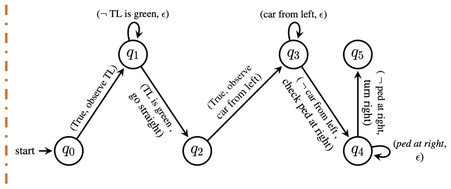
System Modeling



Automaton Construction

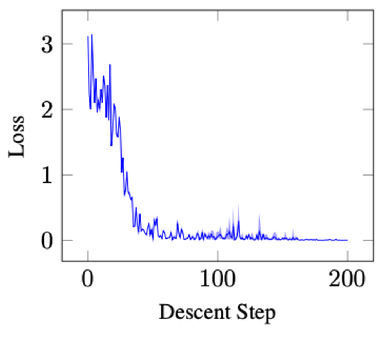
1. Look straight ahead and watch for traffic light.
2. If the traffic light turns green, start moving forward.
3. As you approach the intersection, look to your left for oncoming traffic.
4. If there is no traffic coming from your left, check pedestrians on your right.
5. If it is safe, turn your vehicle right.

```
1. <observe traffic light>.
2. <if> <green traffic light>, <go straight>.
3. <observe car from left>.
4. <if> <no car from left>, <check pedestrian
     at right>.
5. <if> <no pedestrian at right>, <turn
     right>.
```

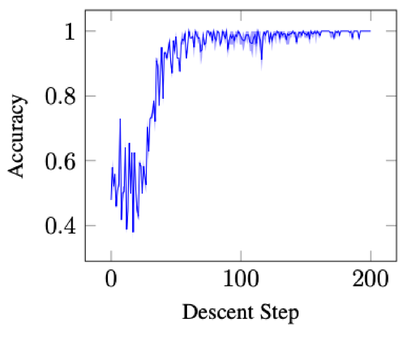


Quantitative Results

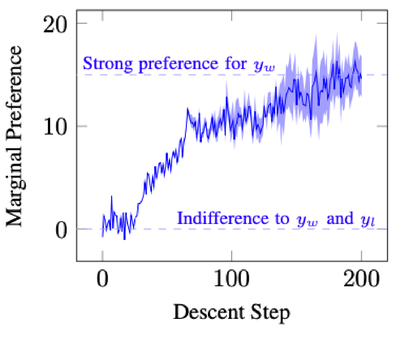
Training Loss



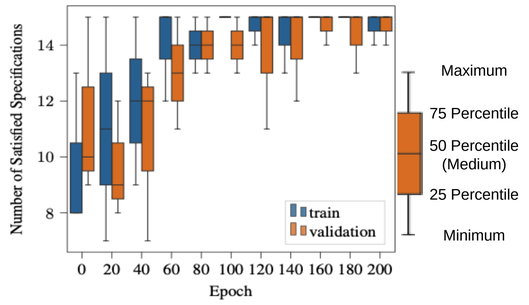
How likely the LLM outputs the preferred response (y_w) rather than the unpreferred response (y_l).



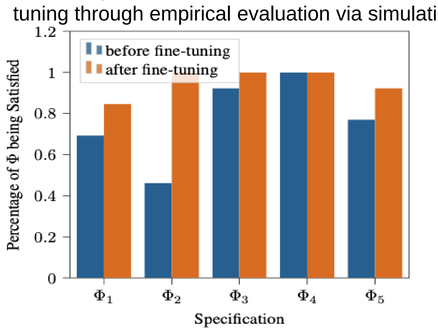
How strongly the LLM favors the preferred response (y_w) rather than the unpreferred response (y_l).



Number of specifications satisfied during fine-tuning via formal verification feedback.



Number of specifications satisfied before and after fine-tuning through empirical evaluation via simulation.



Contact: yunhaoyang234@utexas.edu |