Colin Shea-Blymyer & Houssam Abbas
EECS, Oregon State University

sheablyc@oregonstate.edu
https://web.engr.oregonstate.edu/~sheablyc/
http://www.houssamabbas.com/

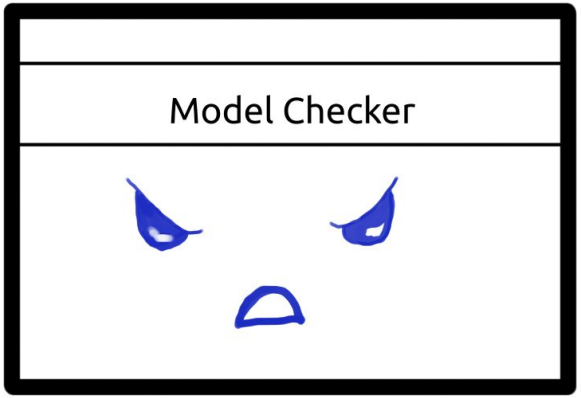
FoMo: Formula and Model Generation

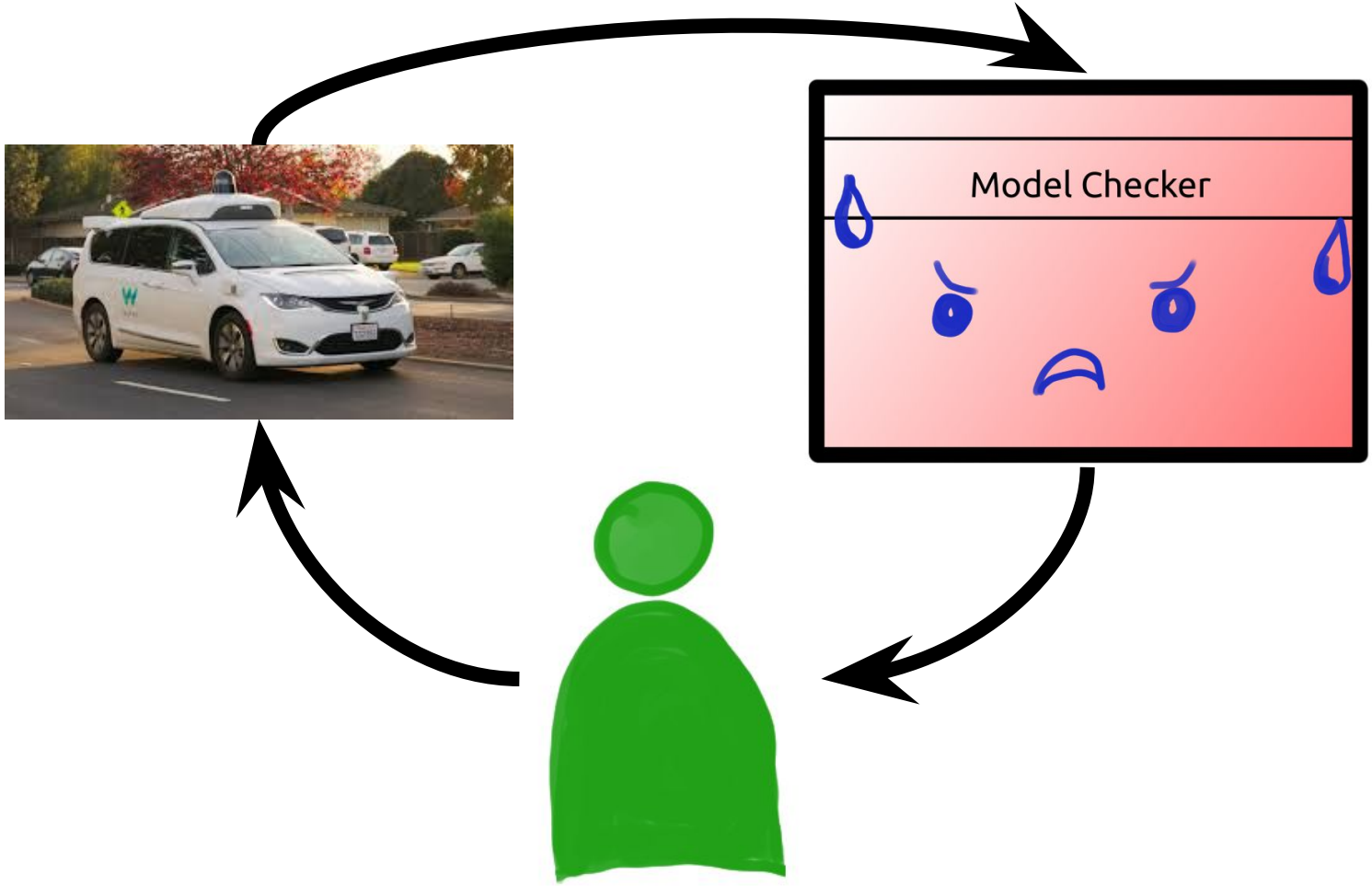
for Learning-Based Formal Methods

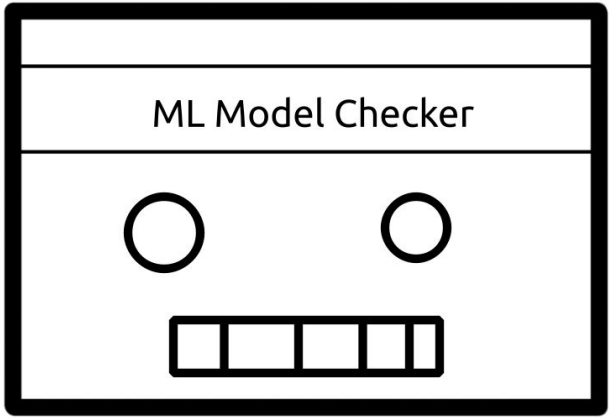
Oregon State University

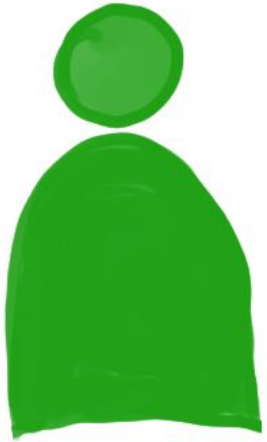
Cyber-Physical Systems are Safety Critical

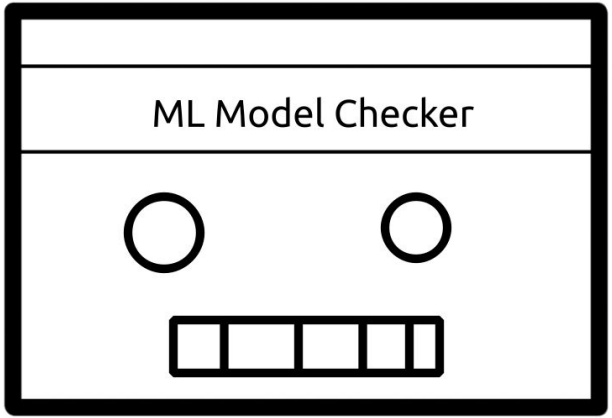
- Autonomous drone deliveries
- Robotic nursing assistants
- Self driving cars

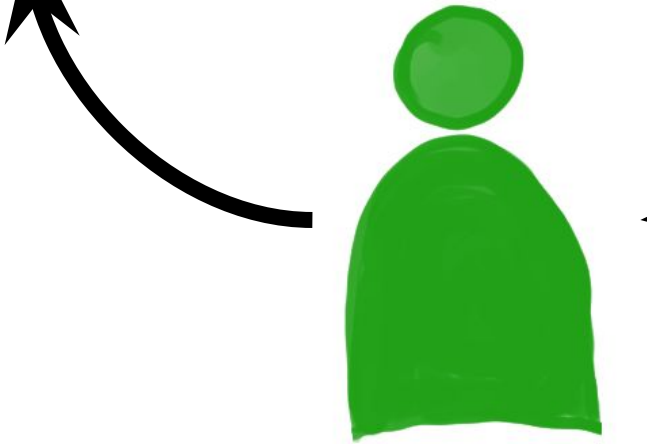

Model Checker

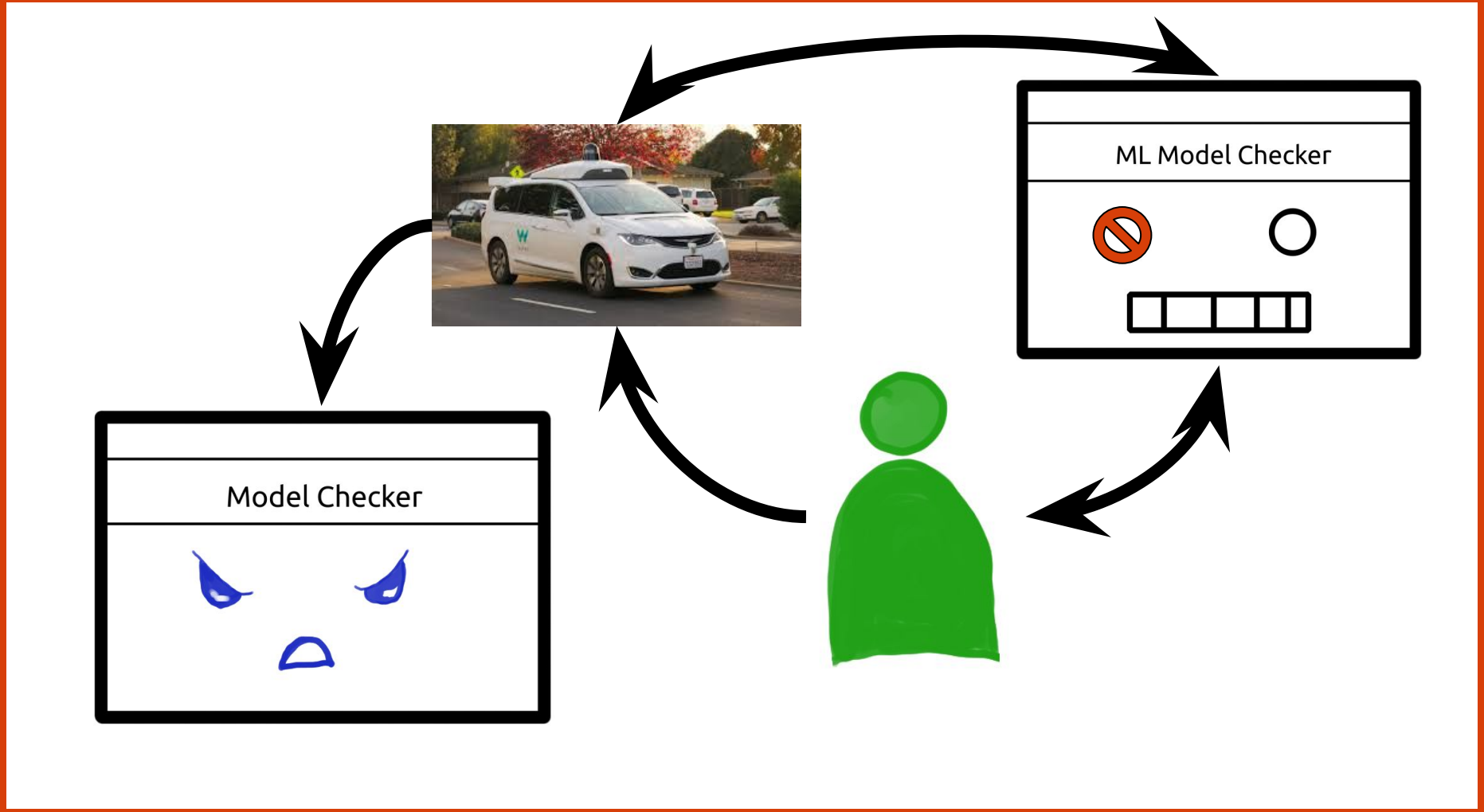

Model Checker

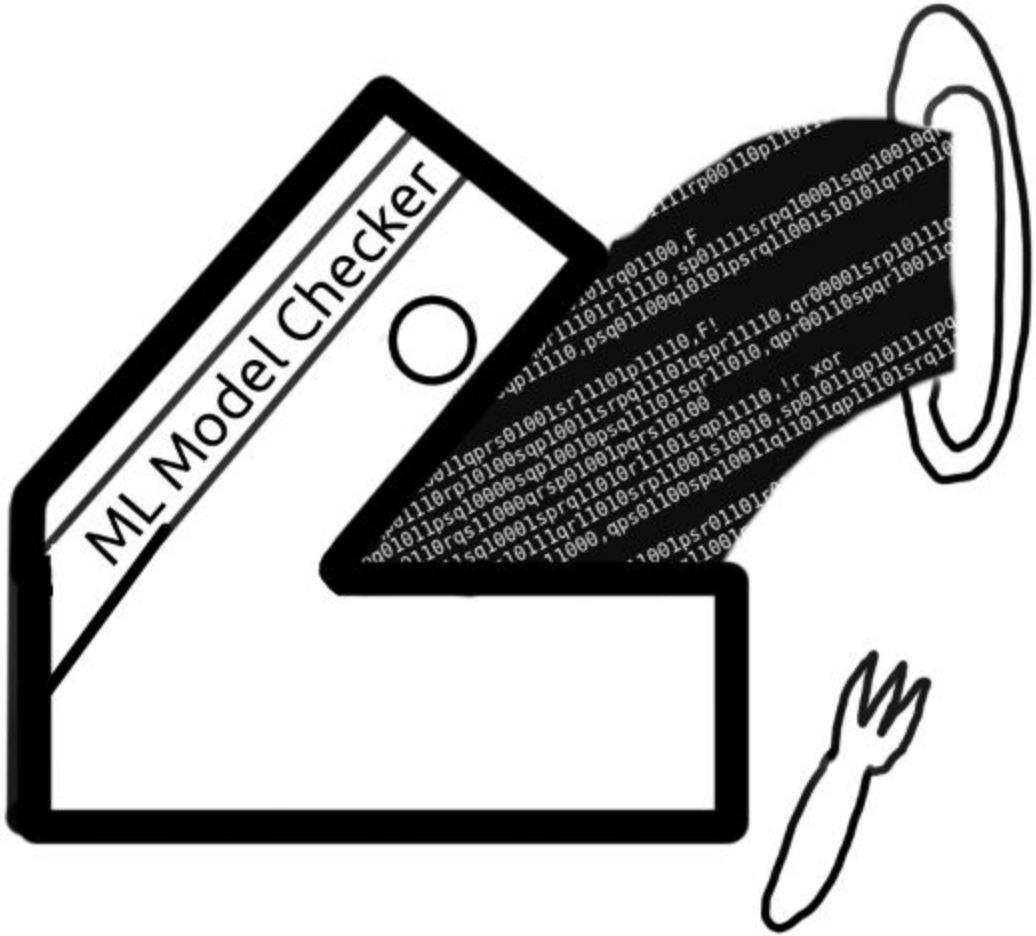

ML Model Checker



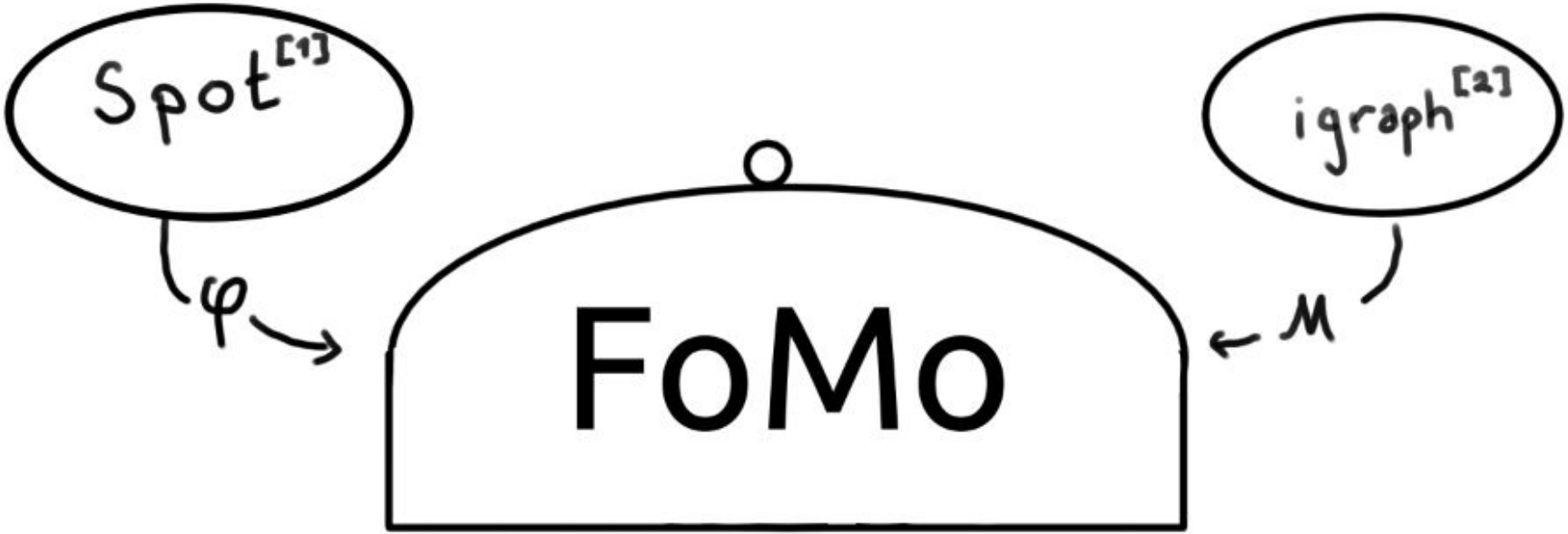

ML Model Checker




ML Model Checker
Model Checker

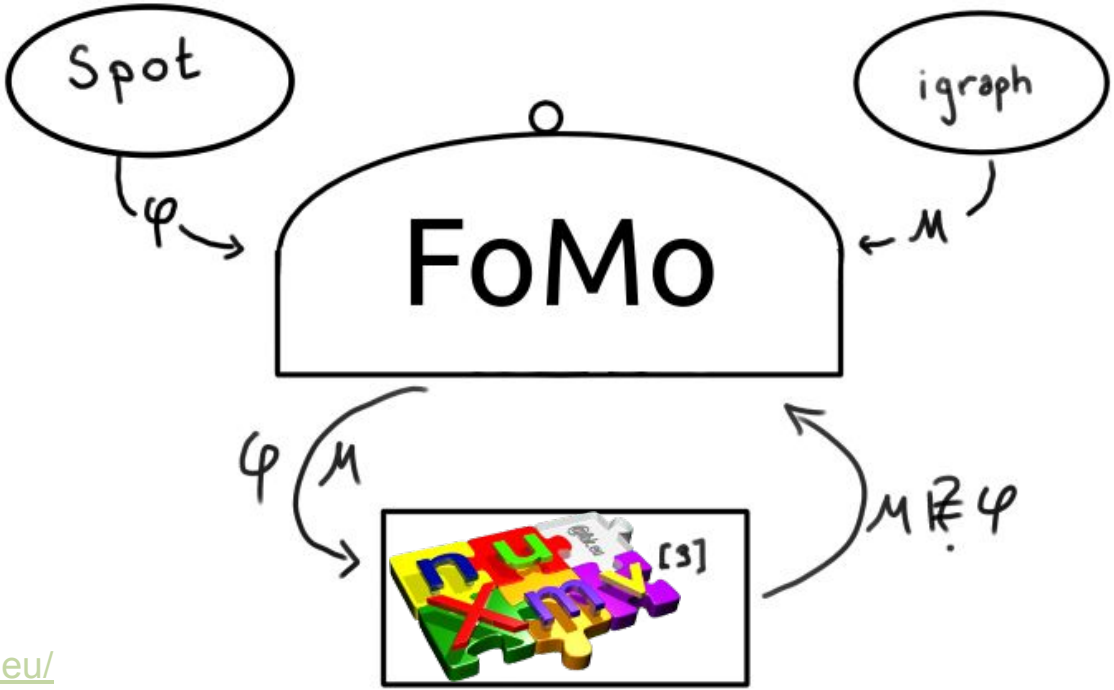

ML Model Checker

FoMo Architecture



[1] https://spot.lrde.epita.fr/
[2] https://igraph.org/

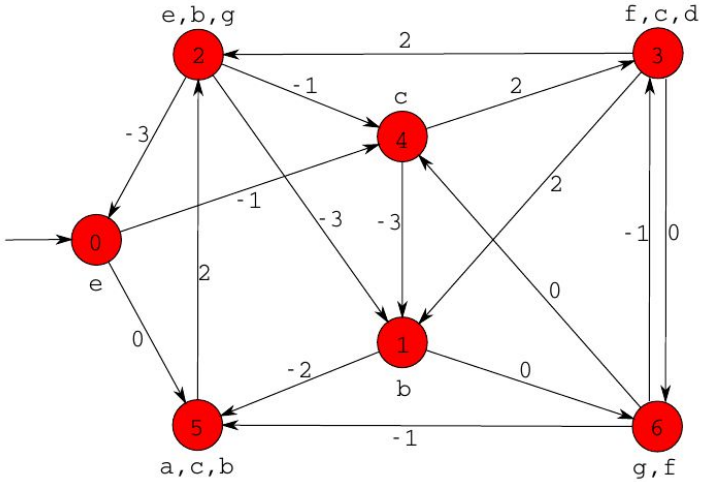
FoMo Architecture



[3] https://nuxmv.fbk.eu/

Generating Models

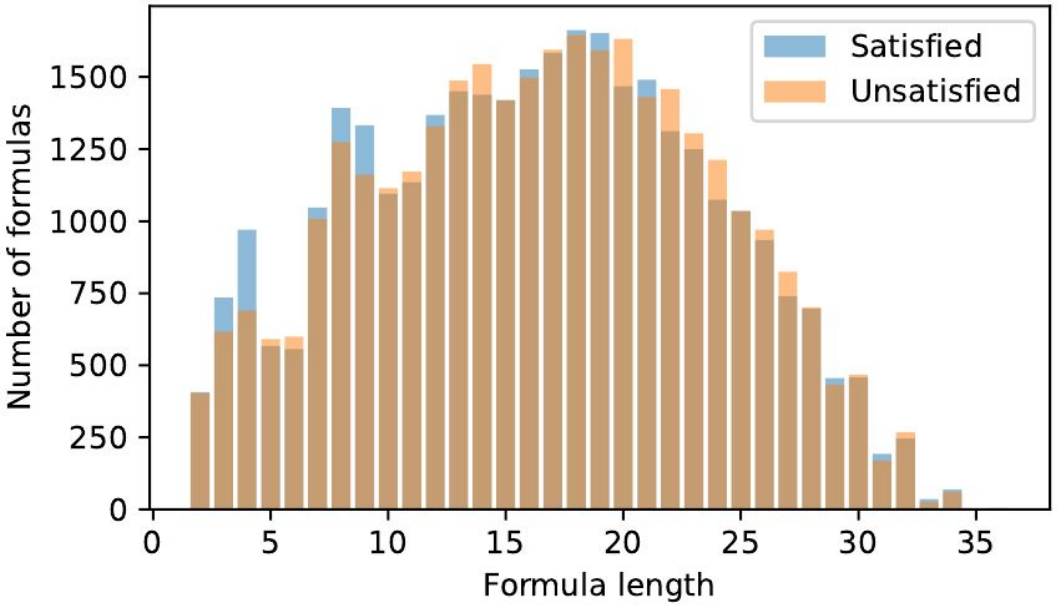
```
Mo =
generate_automaton(generator=Graph.Erdos_Renyi,
symbols=['a', 'b', 'c', 'd', 'e', 'f', 'g'],
max_symbols=2, min_weight=-3, max_weight=3,
n=15, m=3, directed=True)
```

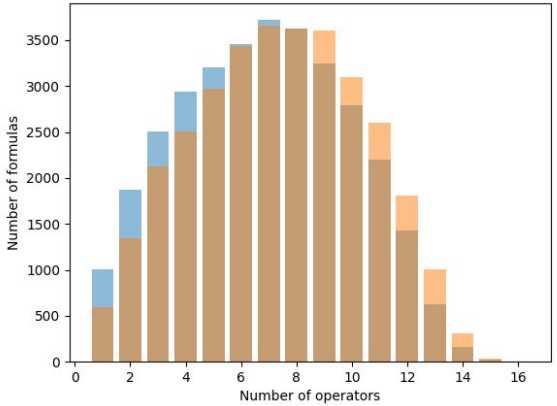


Formula Generation

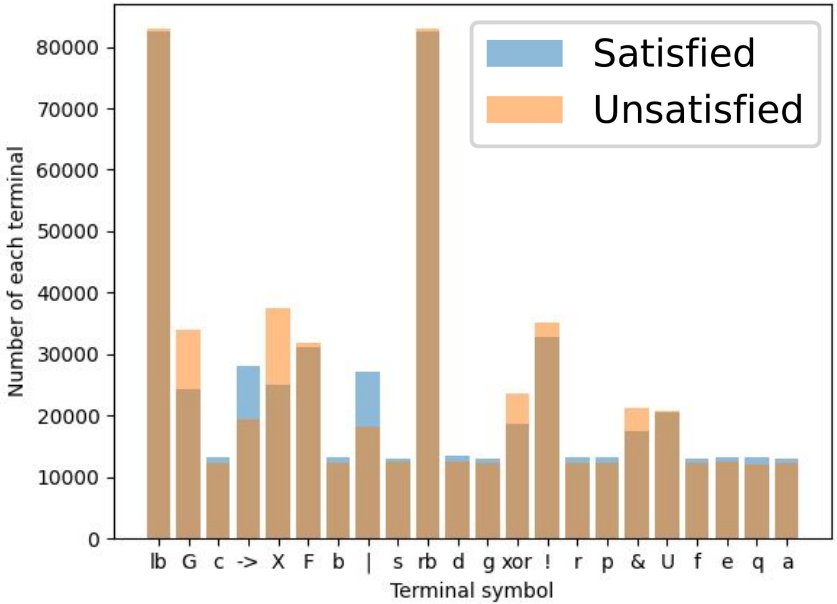
```
Fo =
generate_formula(automaton=Mo,
grammar=['a', 'b', 'c', 'd', 'e', 'f', 'g'],
max_formula_length=7)
```

Formula Distribution

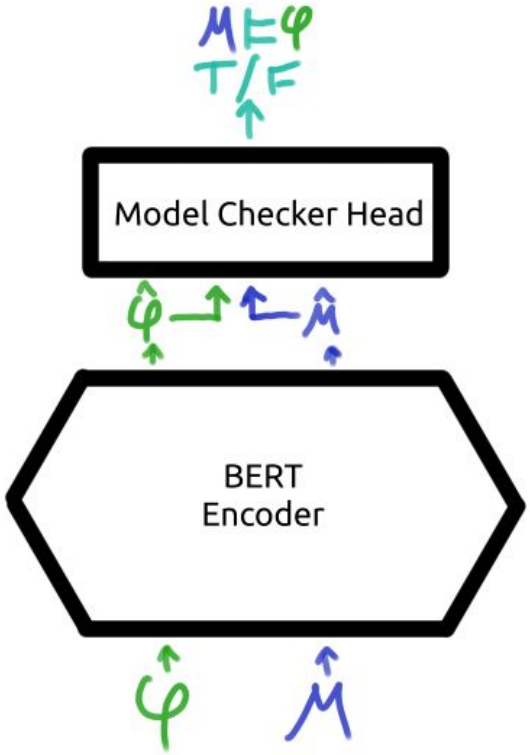




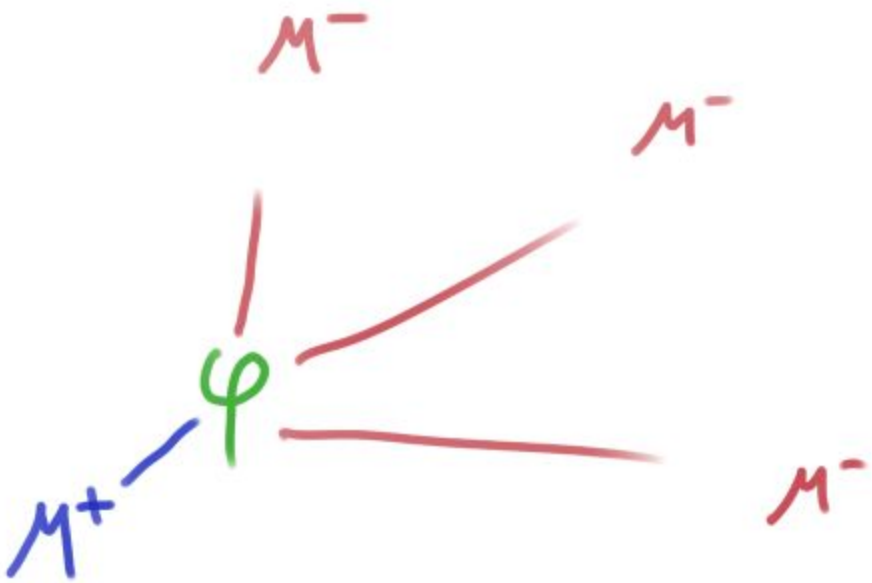
Formula Distribution



Training a Model Checker



Supervised Contrastive Loss



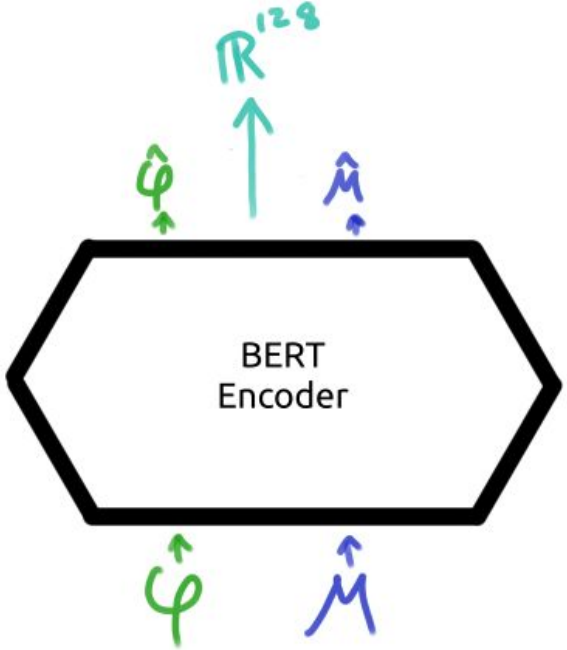
Data

- 130,000 entries
- Erdős–Rényi graphs
  - 10 states
  - 4 of 11 AP per state
- Formulas up to parse tree size 7

Encoder

Trained for 64 epochs

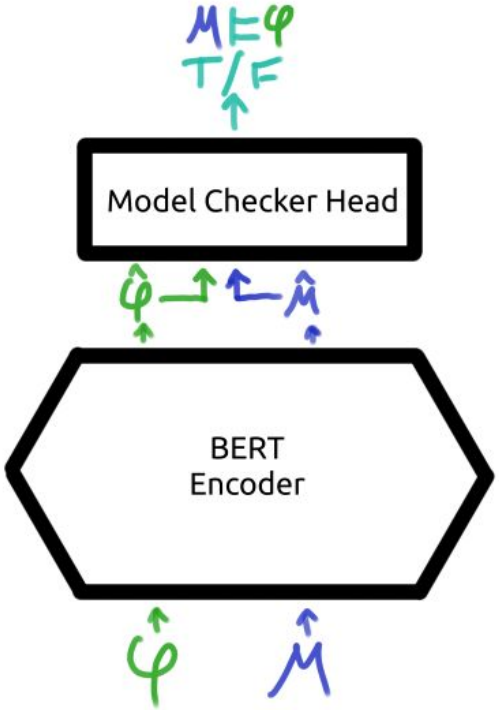
Hidden dimension of 128



Model Checking Head

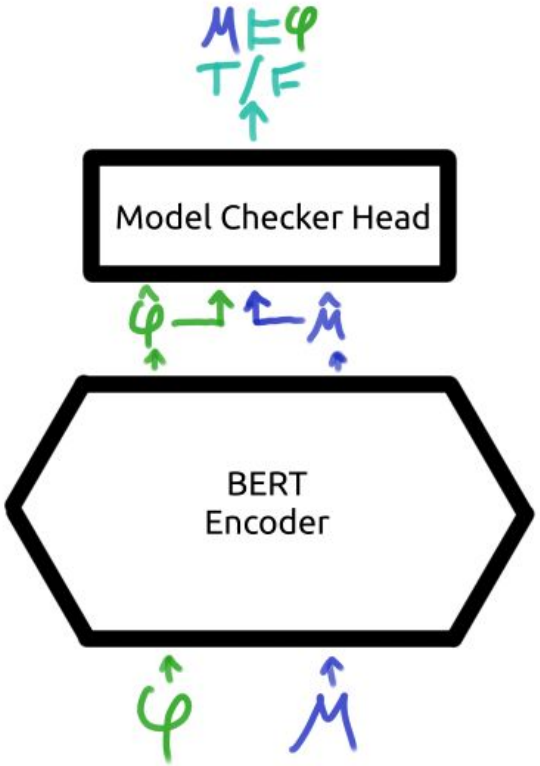
Cross entropy loss

Trained for 10 epochs



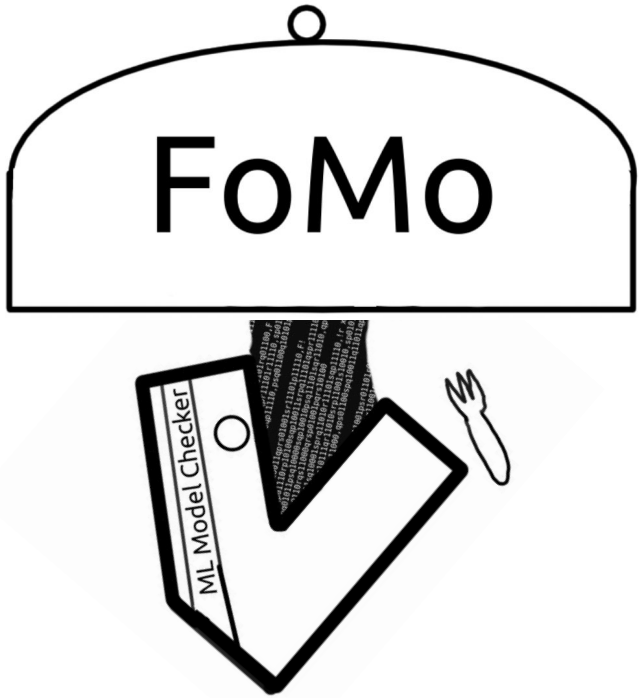
Results

- Validated on Barabasi distribution
- Accuracy of 69%



Conclusion

- Generate data for ML-based verification
- Extend FoMo to probabilistic models and languages
- Refining Deep-Model-Checking



https://github.com/sabotagelab/FoMo

Outline

~ 23 slides; ~20 minutes

- 1-2 slide(s) on why verifying (large) systems matters
- 2 slides on the utility of learning-based methods
  - Scaling up verification in non-critical steps of design
  - Providing non-combinatorial search space
- Overview of tool architecture
- Model generation example
- Formula generation example
- System trace generation example
- 3 slides data distribution
- How can we use all this data? Neural network intro
- Contrastive data generation example
- 3 slides neural network implementation and results