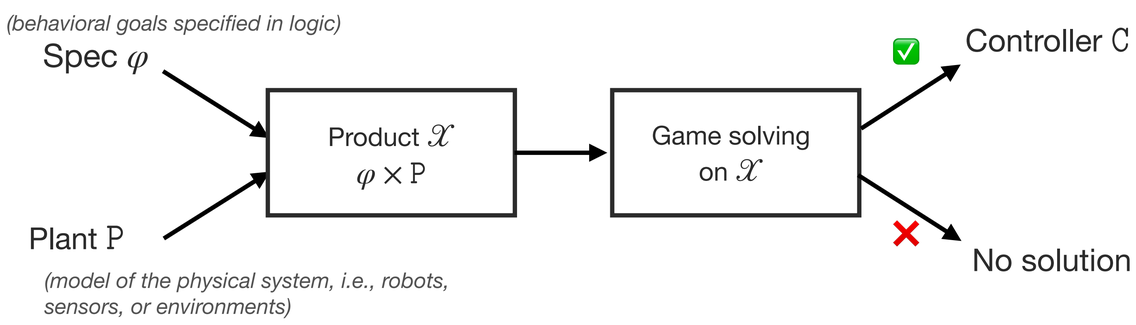
Universal Safety Controllers with Learned Prophecies

Bernd Finkbeiner, Niklas Metzger, Satya Prakash Nayak, Anne-Kathrin Schmuck

Controller Synthesis

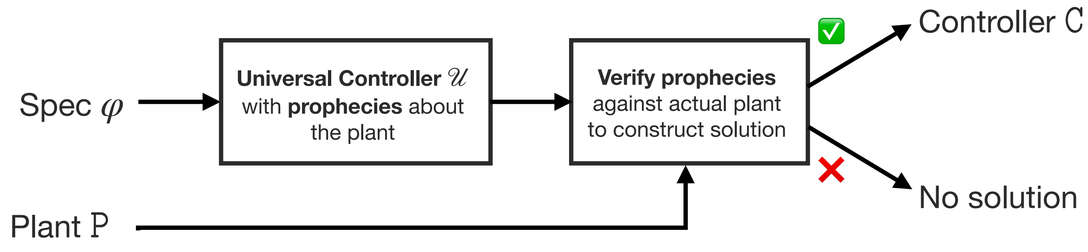


- Controller tied to a **single** plant
- Size of plant **dominates** the approach
- **Full** state-space exploration of the plant

→

- **No** generalization
- **No** scalability
- **No** explainability

Universal Safety Controllers



- **Universal:** provides controller for all plants
- Decisions are guided by **prophecies**, i.e., assumptions on the plant
- **Avoids** state explosion whenever possible

→

- **Strong** generalization
- **More** scalable
- **More** explainable

From Complex Tree Automata to Simple CTL Formulas

Existing Method (Unicon)[1]

- Prophecy construction via **tree automata**
- **Hard** to verify
- **Too complex** to understand

→

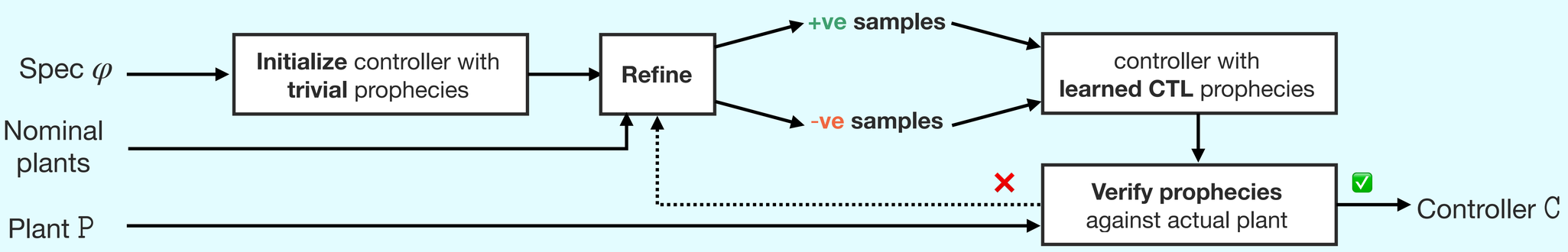
Our Method (UCLearn)

- Prophecy approximation **via CTL learning**[1]
- Refinement via **parity game solving**[3]
- **Polynomial-time** verification
- **Human-readable** CTL formulas

→

Advantages

- **Generalizes** to similar plants
- **Highly** scalable
- **Explainability**

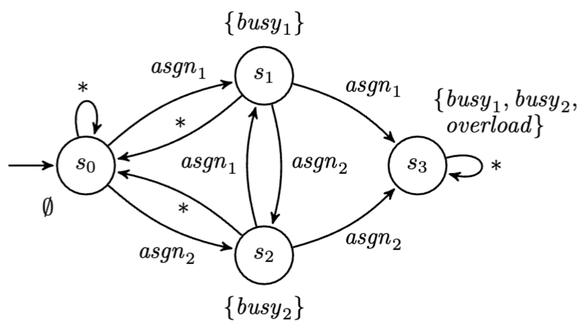


A Load Balancer Example

The controller ***assigns tasks*** to 2 CPUs that can be ***busy,*** free, or ***overloaded***

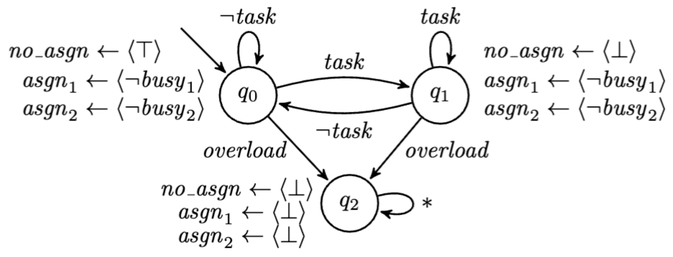
Spec: assign each task to a CPU without being overloaded

$$\varphi = \square(task \rightarrow \bigcirc(asgn_1 \vee asgn_2)) \wedge \square \neg overload$$



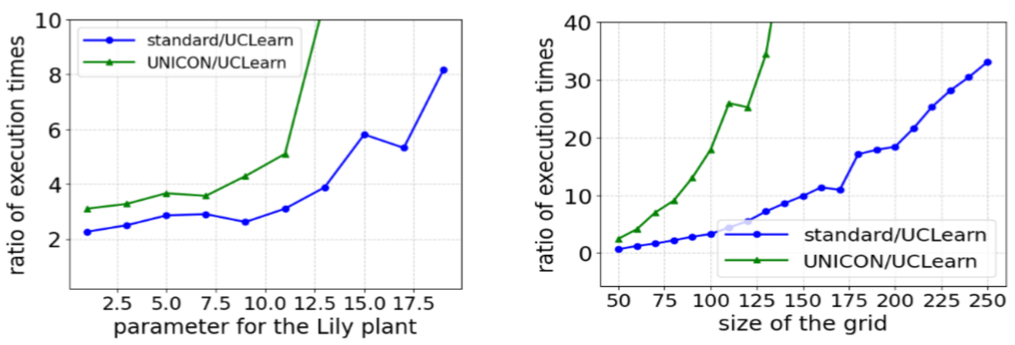
One nominal plant

each of the processors is busy for exactly one time-step once assigned a task

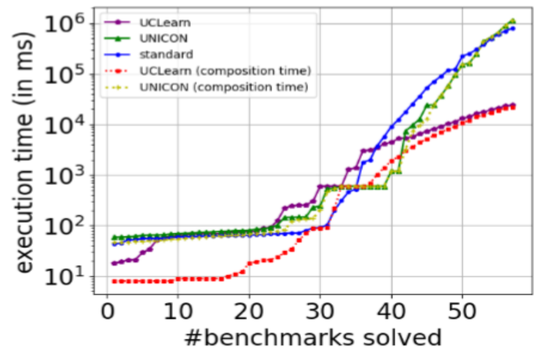


Controller with learned CTL prophecies

Experimental Results



- Learned from **a single nominal plant** with parameter **2**
- **Small** and concise CTL formulas (with size **at most 4**)
- Up to **40x faster** than existing methods



[1] Finkbeiner et al. Synthesis of universal safety controller, TACAS 2025
[2] Bordais et al. Learning branching-time properties in CTL and ATL via constraint solving, FM 2024
[3] van Dijk. Oink: An Implementation and Evaluation of Modern Parity Game Solvers, TACAS 2018.

CISPA HELMHOLTZ-ZENTRUM FÜR INFORMATIONSSICHERHEIT
MAX PLANCK INSTITUTE FOR SOFTWARE SYSTEMS