PROACTIVE CONTEXT-FORECASTED SAFETY CONSTRAINTS FOR NONSTATIONARY REINFORCEMENT LEARNING

Tim Tomashevskiy
Department of Computing and Software · McMaster University

McMaster University

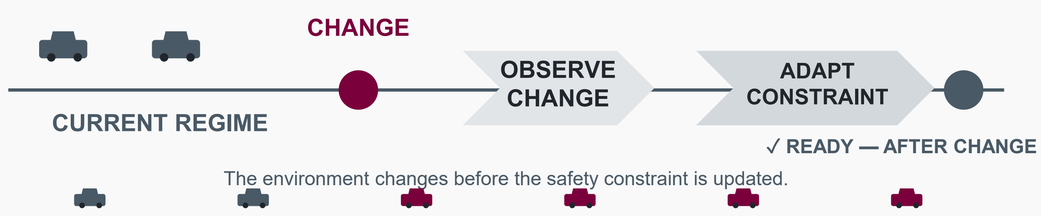
SCAN FOR PAPER

WHY PROACTIVE SAFETY?

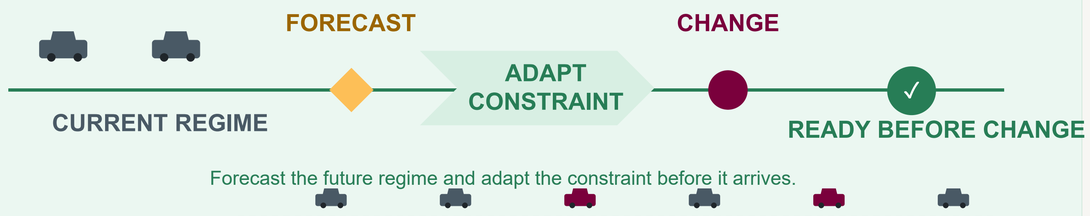
Safety constraints designed for one operating regime can become outdated or overly conservative as the environment changes.

CONCEPTUAL COMPARISON

TYPICAL REACTIVE ADAPTATION



PROACTIVE — OUR APPROACH



SAFETY CONSTRAINTS SHOULD ANTICIPATE CHANGE — NOT MERELY REACT TO IT.

HOW DOES IT WORK?

OBSERVE (from experience) → **UNDERSTAND** (infer the regime) → **ANTICIPATE** (forecast risk) → **ACT** (filter before execution)

RECENT TRANSITIONS
state · action · next state
Observe how the environment is currently behaving



→

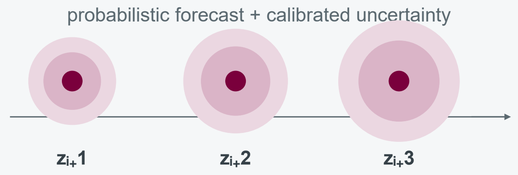
CURRENT LATENT CONTEXT
context discovery



z_i = inferred episode regime

→

FORECAST FUTURE EPISODE CONTEXTS
probabilistic forecast + calibrated uncertainty



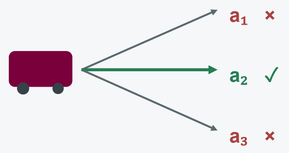
uncertainty grows with the forecast horizon

↓

TAIL-SAFE FUTURE CONSTRAINTS c'_{t+1} c'_{t+2} · · · c'_{t+n}
future-step safety margins derived from the anticipated regime

→

MPC-STYLE SAFETY FILTER
evaluate candidate actions against forecast constraints



→ **✓ SAFE ACTION**
execute only feasible action

Reject actions predicted to violate future safety constraints.

Infer current context → forecast future context with calibrated uncertainty → anticipate future safety constraints → filter actions before execution.

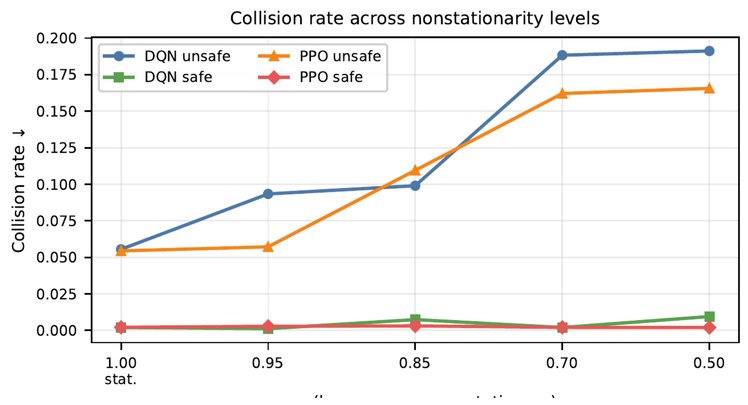
EXPERIMENTAL EVIDENCE

DOES IT WORK?

98.9% COLLISION REDUCTION

18.83% → 0.20%

DQN · strong nonstationarity · p_{stay} = 0.70

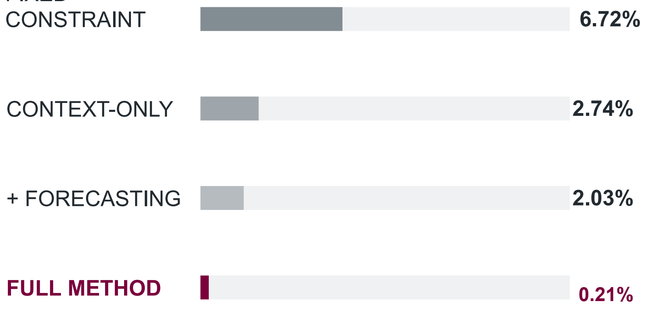


PPO: 16.21% → 0.21%
Same strong-nonstationarity setting

>90% reduction across every evaluated setting
usually >97%

Unsafe policies deteriorate as context switching increases; the safety-enabled curves remain near zero.

WHICH COMPONENTS MATTER?



17.52 → 6.72 → 2.74 → 2.03 → 0.21%
No safety → Fixed → Context → Forecast → Full

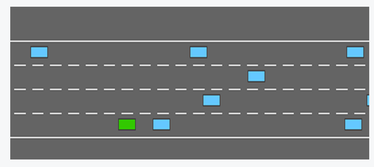
Static safety helps; context dependence and forecasting progressively reduce collisions further.

Component analysis at p_{stay} = 0.70; aggregated across DQN and PPO.

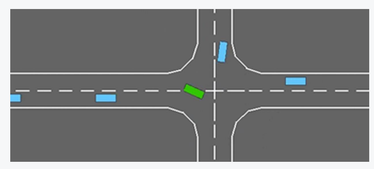
FULL METHOD · 0.21% COLLISIONS

HELD-OUT LAYOUT STRESS TEST

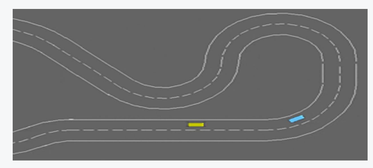
Representative HighwayEnv environment views



HIGHWAY-v0
97.1% ↓
collision reduction



INTERSECTION-v0
92.0% ↓
collision reduction



RACETRACK-v0
95.7% ↓
collision reduction

Collision reductions persist across all evaluated held-out layouts.

Percentages aggregate DQN/PPO across evaluated nonstationarity levels.

Representative views: HighwayEnv / Farama Foundation

WHAT SAFETY CLAIM DO WE MAKE?

CONDITIONAL HORIZON-LEVEL SAFETY

$$\Pr(\exists\, k \le n : d_{t+k} < \varepsilon) \le \alpha + nq$$

Probability of a clearance violation within the forecast horizon is bounded by α + nq.

CALIBRATED CONTEXT COVERAGE $1 - \alpha$ + **TAIL-SAFE CLEARANCE** PREDICTION q + **CONTROL-AWARE** MPC FILTERING

HORIZON-LEVEL SAFETY BOUND $\alpha + nq$

Conditional probabilistic guarantee over the forecast horizon — not unconditional hard safety.
Requires calibrated context coverage, tail-prediction coverage, and MPC feasibility.

TAKE-HOME

Proactive context forecasting enables safety constraints to adapt before anticipated environmental change.

FORECAST THE CONTEXT.
ADAPT THE CONSTRAINT.
ACT BEFORE THE RISK ARRIVES.

tomashet@mcmaster.ca

SCAN FOR PAPER

CoLLAs 2026