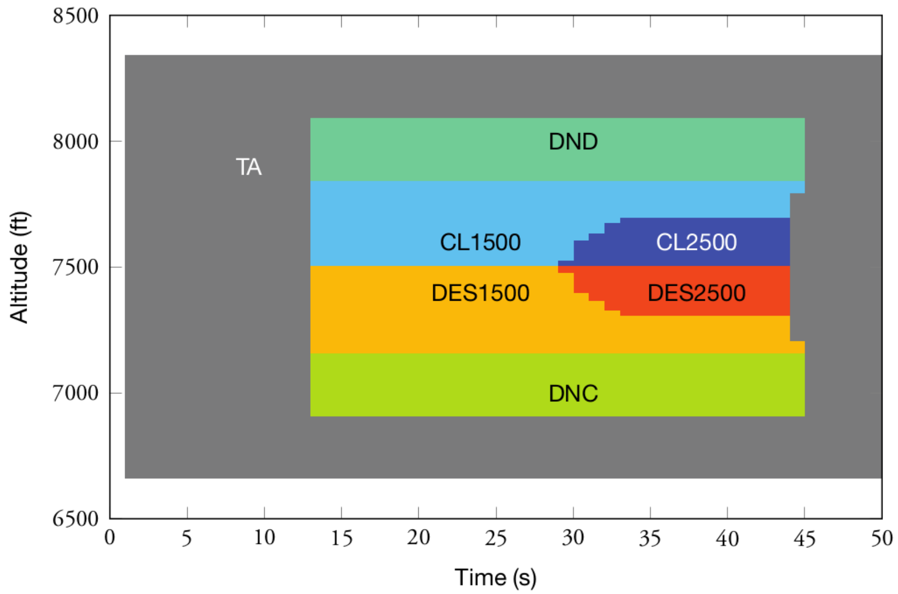
POMDPs in Aerospace and Autonomy

1. ACASX
2. Autonomous Driving

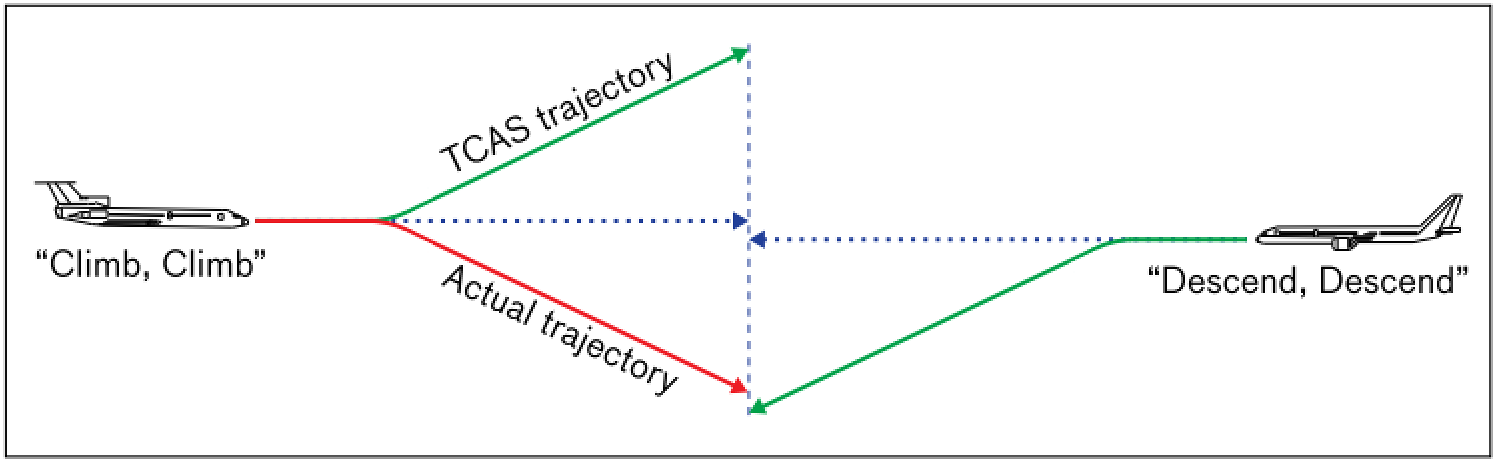
1. ACASX

TCAS

```
IF (ITF.A LT G.ZTHR)
  THEN IF(ABS
  (ITF.VMD) LT
  G.ZTHR)
    THEN SET ZHIT;
    ELSE CLEAR
```



(a) TCAS.



ACASX POMDP

State variables

- h, the altitude of the intruder relative to the own aircraft,
- $\dot{h}_0$, the vertical rate of the own aircraft,
- $\dot{h}_1$, the vertical rate of the intruder aircraft,
- τ, the time to potential collision,
- s_{adv}, the current advisory, and
- s_{res}, whether the pilot is responding to the advisory.

Transitions

$$\begin{bmatrix} h \\ \dot{h}_0 \\ \dot{h}_1 \\ \tau \\ s_{\text{adv}} \\ s_{\text{res}} \end{bmatrix} \leftarrow \begin{bmatrix} h + \dot{h}_1(\Delta t) + \frac{1}{2}\ddot{h}_1(\Delta t)^2 - \dot{h}_0(\Delta t) - \frac{1}{2}\ddot{h}_0(\Delta t)^2 \\ \dot{h}_0 + \ddot{h}_0(\Delta t) \\ \dot{h}_1 + \ddot{h}_1(\Delta t) \\ \tau - 1 \\ a \\ s'_{\text{res}} \end{bmatrix}. \tag{10.2}$$

$$P(s'_{\text{res}} = \text{true} \mid s_{\text{adv}}, s_{\text{res}}, a) = \begin{cases} 1 & \text{if } a = \text{COC} \\ 1 & \text{if } s_{\text{res}} = \text{true and } s_{\text{adv}} = a \\ 1/(1+5) & \text{if } s_{\text{adv}} = \text{COC and } a \neq \text{COC} \\ 1/(1+5) & \text{if } s_{\text{adv}} \text{ and } a \text{ are opposite sense} \\ 1/(1+3) & \text{if } s_{\text{adv}} \text{ and } a \text{ are same sense} \end{cases}. \tag{10.1}$$

Actions

- do not climb or descend,
- limit climb or descend to 500, 1000, or 2000 ft/min,
- level off,
- climb or descend at 1500 ft/min,
- increase climb or descend to 2500 ft/min, or
- maintain current vertical rate.

Observations: Noisy and quantized measurements of h, $\dot{h}_0$, $\dot{h}_1$

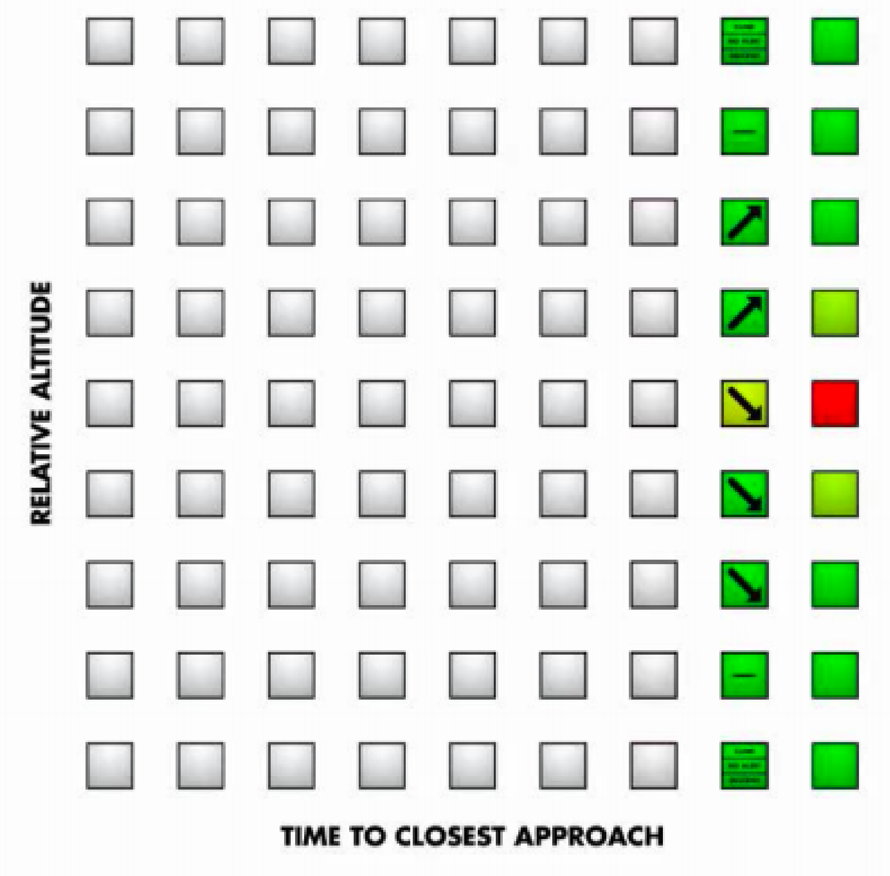
ACASX POMDP: Reward

Reward	Separation (ft)	Closure (ft/min)	Event
-1	≤ 175		$\tau \leq 0$
-1			Maintain advisory with $\dot{h}_0 < 1500$ ft/min
-1			Prohibited advisory transitions
-1			Preventive crossing advisory
-0.1	> 650	< 2000	Corrective advisory
-3×10^{-2}	> 1000	< 4000	Corrective advisory
-1×10^{-2}	> 650	< 2000	Preventive advisory
-1×10^{-2}	> 500		Crossing advisory
-8×10^{-3}			Reversal
-5×10^{-3}			Strengthening
-1×10^{-3}			Weakening
-1.5×10^{-3}		> 3000	Non-MVS/LOLO
-2.3×10^{-3}		< 3000	Any advisory
-5×10^{-4}		> 3000	MVS/LOLO
$-4 \times 10^{-4} \times \Delta\dot{h}$			Crossing advisory when $\lvert\dot{h}_0\rvert > 500$ ft/min and $\dot{h}_0$ is in opposite direction of advisory
-4×10^{-4}			Maintain
-1×10^{-4}			MVS/LOLO
$-3 \times 10^{-5} \times \Delta\dot{h}$			Any advisory
-1×10^{-5}			Corrective advisory
1×10^{-9}			COC

ACASX: QMDP Solution

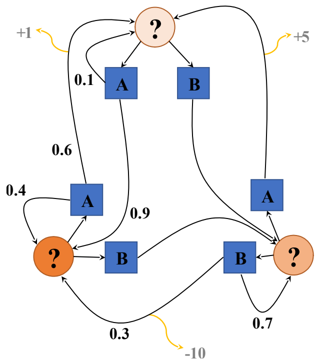
Because the state space is continuous, the local approximation value iteration algorithm (Section 4.5.1) is used to compute U^*. In particular, ACAS X uses multilinear interpolation over a grid-based discretization of the state space. The h variable is discretized into 33 points over the range ±4000 ft with finer discretization near 0 ft. The vertical rate variables $\dot{h}_0$ and $\dot{h}_1$ are discretized into 25 points each between ±10,000 ft/min with finer discretization near 0 ft/min. The time to potential collision variable τ is discretized at 1 s from 0 to 40 s.

The structure of the problem is such that only a single sweep of Gauss-Seidel value iteration (Section 4.2.6) is required. Because states with $\tau = k$ depend only on the states with $\tau = k - 1$, ordering the sweep of the states by increasing τ value results in the optimal value function. Although there are more than 26 million vertices, the process requires only a few minutes on a modern workstation. The state-action values $Q^*(s, a)$ produced through dynamic programming are saved in a lookup table. Recent research has explored methods for reducing the size of the table for use on airborne equipment with limited memory capacity [9].

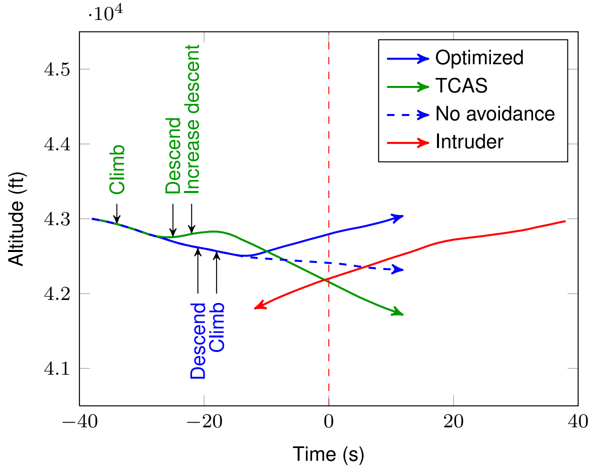


ACAS X

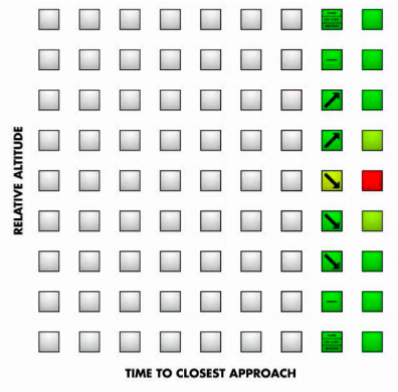
POMDP Models



+ julia =



Optimization



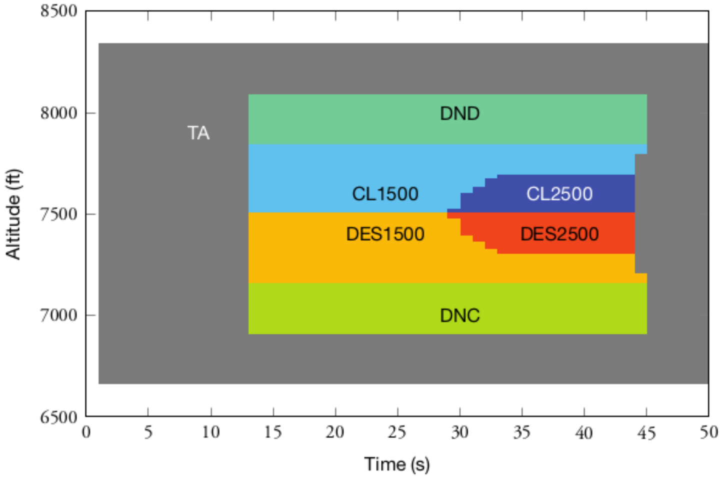
Specification

Algorithm 11 ProximityEstimation

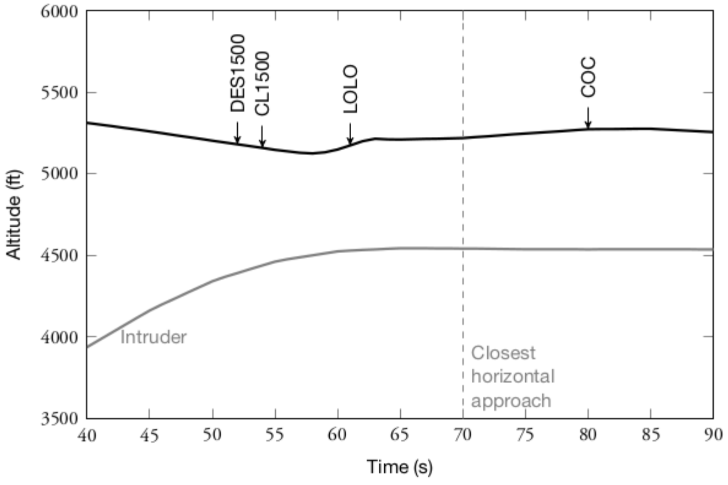
```
function ProximityEstimation( z_own_ave::R, z_int_ave::R, r_ground_int::R, valid_bearing_int::Bool,
                              belief_vert_int::Vector{IntruderVerticalBelief(p. 102)}, quant_int::Z )
    D_proximity_range_threshold::R = params().display.proximity_range_threshold
    H_proximity_altitude_threshold = params().display.proximity_altitude_threshold
    is_proximate::Bool = false
    z_rel::R = abs( z_own_ave - z_int_ave )
    if (r_ground_int <= D_proximity_range_threshold) && (z_rel < H_proximity_altitude_threshold)
        if !IsNARS(p. 95)( belief_vert_int, quant_int ) || valid_bearing_int
            is_proximate = true
        end
    end
    return is_proximate::Bool
end
```

ACASX

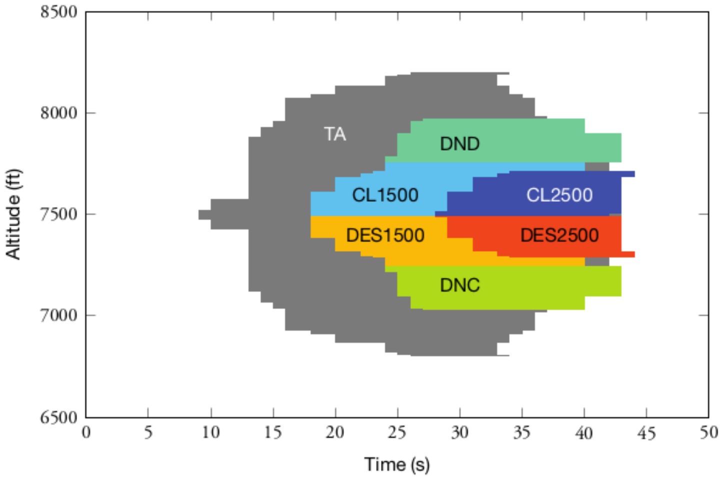
Safer (especially when pilots don't respond) and much fewer advisories.



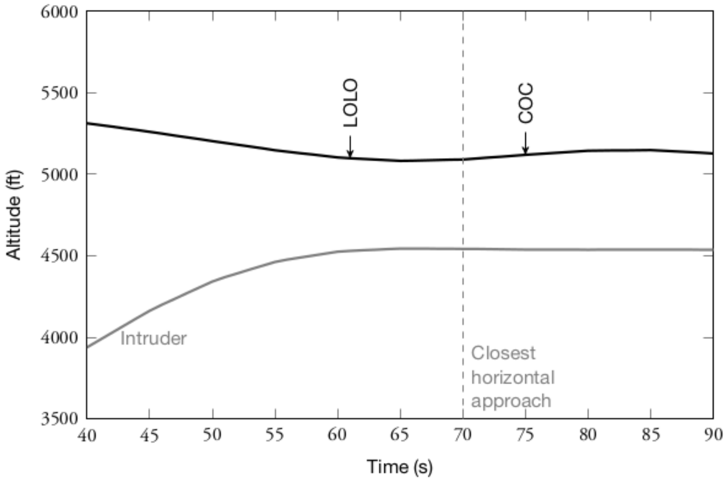
(a) TCAS.



(a) TCAS.



(b) ACAS X.



(b) ACAS X.

2. Autonomous Driving

External States

Position
Velocity
Turn Signals
Conditions

Internal States

Intentions
Disposition
Distraction
Focus

Sadigh, Dorsa, et al. "Information gathering actions over human internal state." *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2016.

Schmerling, Edward, et al. "Multimodal Probabilistic Model-Based Planning for Human-Robot Interaction." *arXiv preprint arXiv:1710.09483* (2017).

Sadigh, Dorsa, et al. "Planning for Autonomous Cars that Leverage Effects on Human Actions." *Robotics: Science and Systems*. 2016.

Tweet by Nitin Gupta
29 April 2018
https://twitter.com/nitguptaa/status/990683818825736192

Human Behavior Model: IDM and MOBIL

$$\ddot{x}_{\text{IDM}} = a\left[1 - \left(\frac{\dot{x}}{\dot{x}_0}\right)^{\delta} - \left(\frac{g^*(\dot{x}, \Delta\dot{x})}{g}\right)^2\right]$$

$$g^*(\dot{x}, \Delta\dot{x}) = g_0 + T\dot{x} + \frac{\dot{x}\Delta\dot{x}}{2\sqrt{ab}}$$

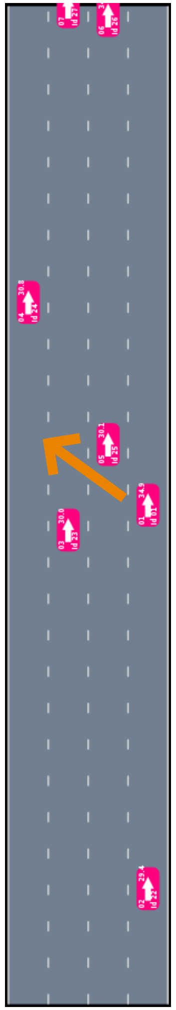
IDM Parameter		Timid	Normal	Aggressive
Desired speed (m/s)	$\dot{x}_0$	27.8	33.3	38.9
Desired time gap (s)	T	2.0	1.5	1.0
Jam distance (m)	g_0	4.0	2.0	0.0
Max acceleration (m/s^2)	a	0.8	1.4	2.0
Desired deceleration (m/s^2)	b	1.0	2.0	3.0
MOBIL Parameter		**Timid**	**Normal**	**Aggressive**
Politeness	p	1.0	0.5	0.0
Safe braking (m/s^2)	b_{safe}	1.0	2.0	3.0
Acceleration threshold (m/s^2)	a_{thr}	0.2	0.1	0.0

M. Treiber, et al., “Congested traffic states in empirical observations and microscopic simulations,” *Physical Review E*, vol. 62, no. 2 (2000).

A. Kesting, et al., “General lane-changing model MOBIL for car-following models,” *Transportation Research Record*, vol. 1999 (2007).

A. Kesting, et al., "Agents for Traffic Simulation." *Multi-Agent Systems: Simulation and Applications*. CRC Press (2009).

POMDP Formulation



Ego physical state

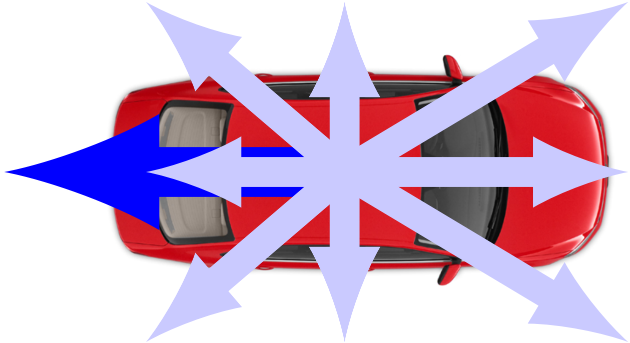
Physical states of other cars

Internal states of other cars

$$s = (x, y, \dot{x}, \{(x_c, y_c, \dot{x}_c, l_c, \theta_c)\}_{c=1}^{n})$$

$$o = \{(x_c, y_c, \dot{x}_c, l_c)\}_{c=1}^{n}$$ ← Physical states of other cars

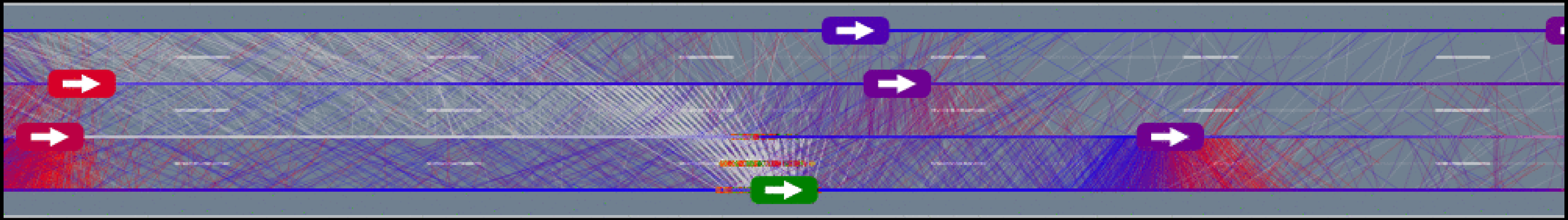
$$a = (\ddot{x}, \dot{y}),\ \ddot{x} \in \{0, \pm 1 \text{ m/s}^2\},\ \dot{y} \in \{0, \pm 0.67 \text{ m/s}\}$$



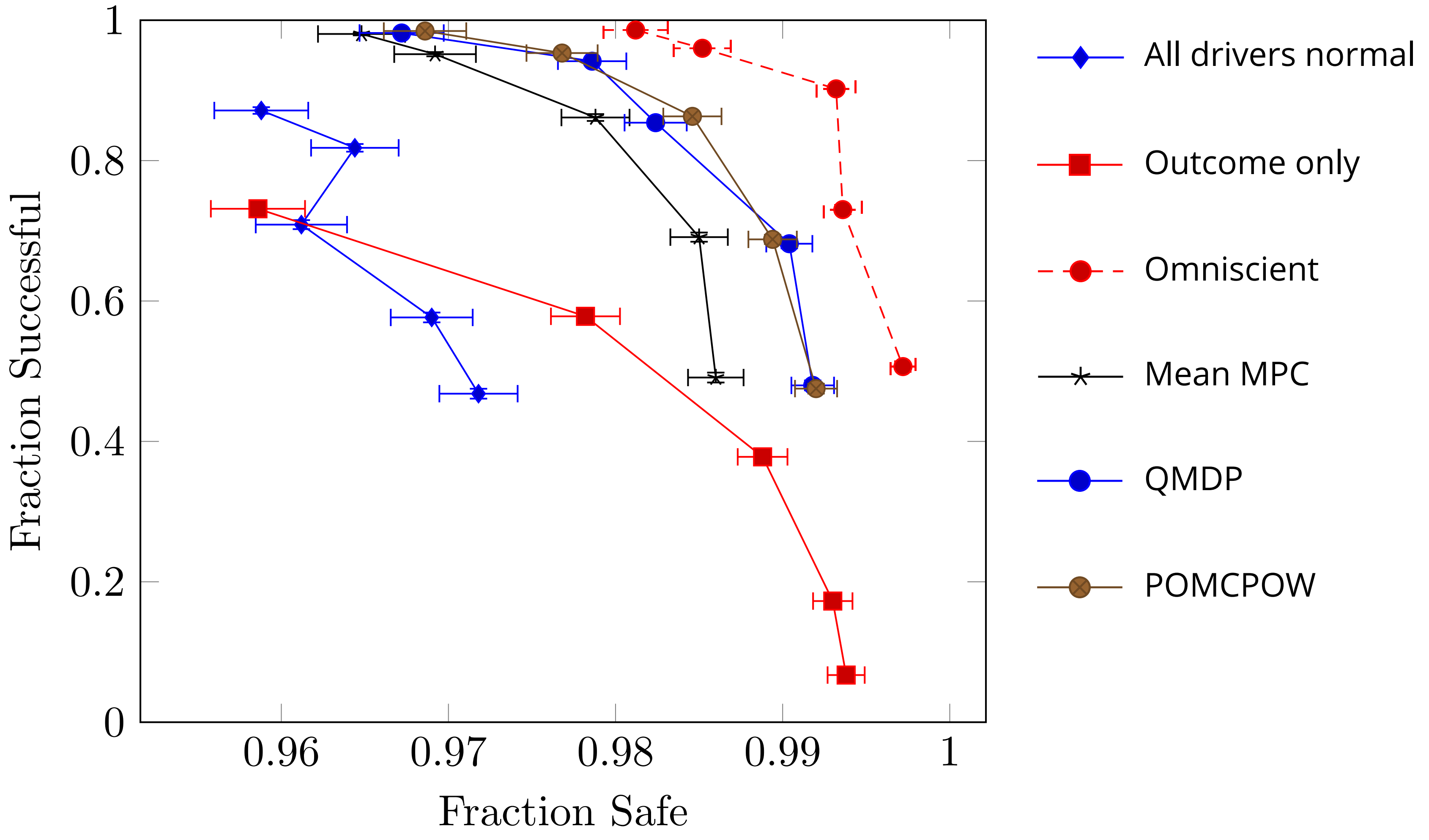
- Actions filtered so they can never cause crashes
- Braking action always available

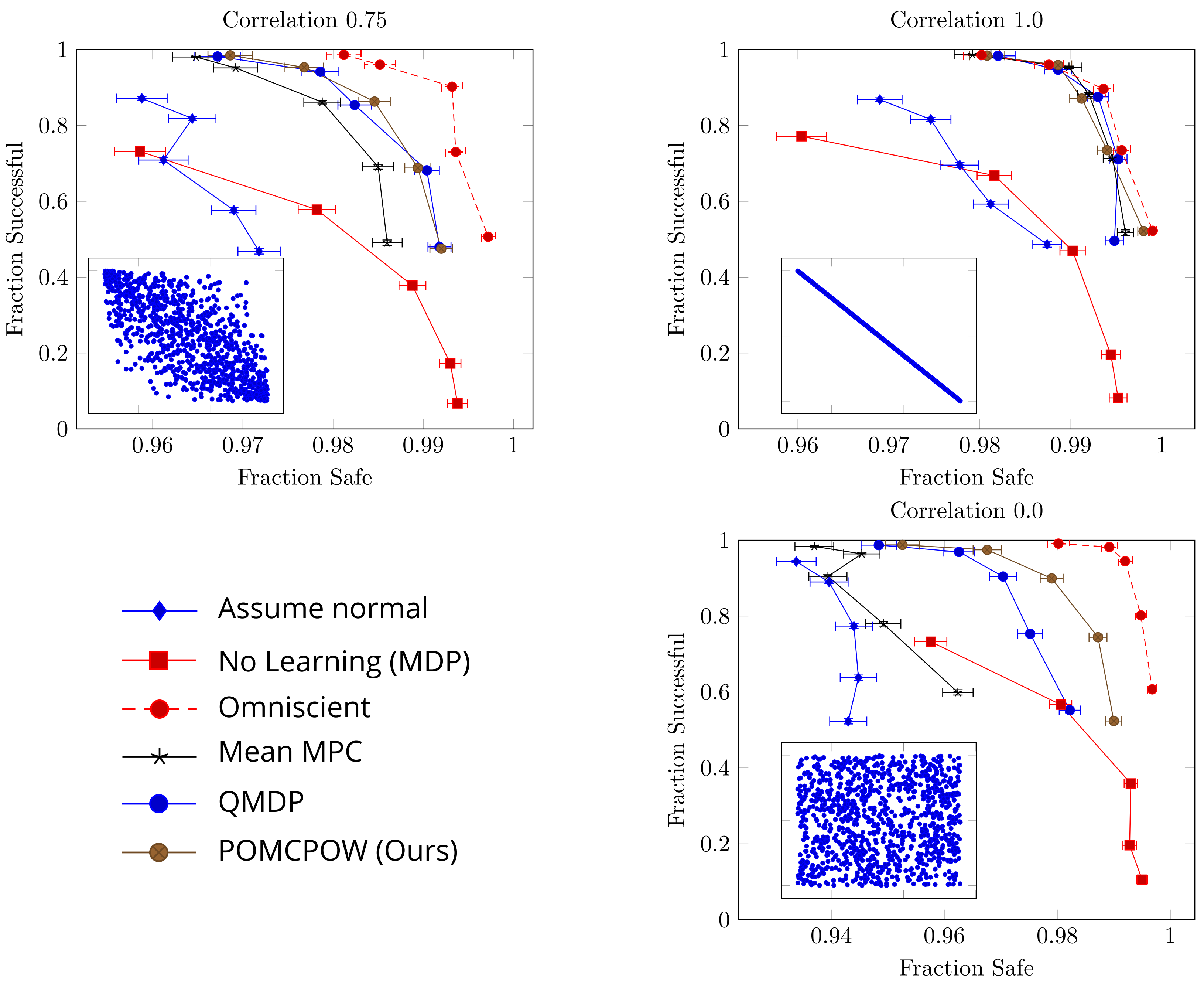
$$R(s, a, s') = \underbrace{\text{in\_goal}(s')}_{\text{Efficiency}} - \lambda \underbrace{\left(\text{any\_hard\_brakes}(s, s') + \text{any\_too\_slow}(s')\right)}_{\text{Safety}}$$

t: 0.15
x: 4.87
vel: 32.89
r: 0.00

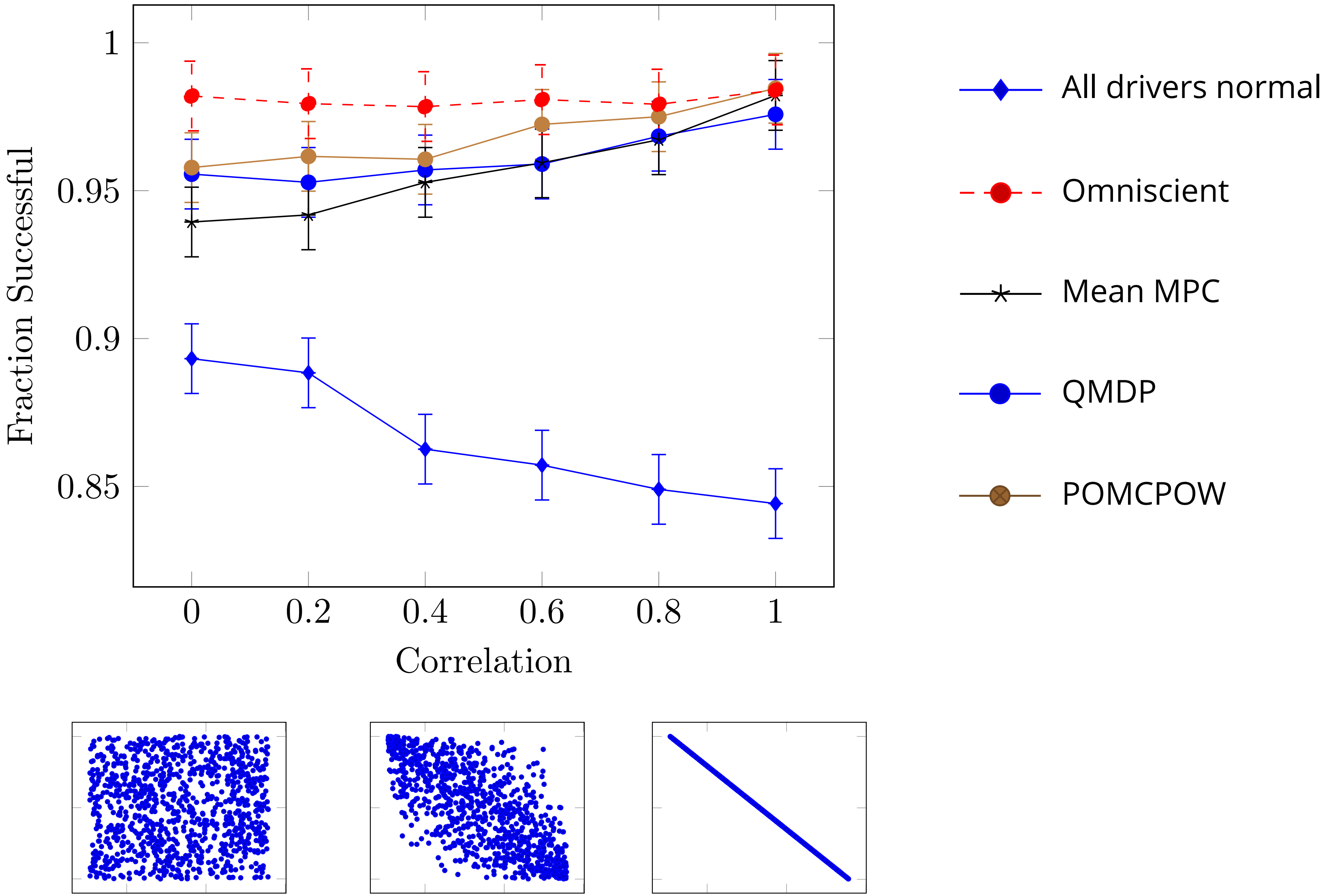


Simulation results





Correlation 0.75
Fraction Successful
Fraction Safe
0 0.2 0.4 0.6 0.8 1
0.96 0.97 0.98 0.99 1
Correlation 1.0
Fraction Successful
Fraction Safe
0 0.2 0.4 0.6 0.8 1
0.96 0.97 0.98 0.99 1
Correlation 0.0
Fraction Successful
Fraction Safe
0 0.2 0.4 0.6 0.8 1
0.94 0.96 0.98 1
Assume normal
No Learning (MDP)
Omniscient
Mean MPC
QMDP
POMCPOW (Ours)



1
0.95
0.9
0.85
Fraction Successful
0
0.2
0.4
0.6
0.8
1
Correlation
All drivers normal
Omniscient
Mean MPC
QMDP
POMCPOW