Not everything is relevant for safety

To prevent a taxi from running out of fuel it is not necessary to know the position of the passenger.

- Constrained MDPs models safety requirements explicitly.
- How to learn without violating the safety constraints?

**Constrained RL**

\[ M = (S, A, P, R, \mu, C, \hat{c}) \]

\[
\max_{\pi} V^\pi(\mu) = E_s \sum_{t=1}^H \sum_{i=1}^{n_i} r_i | \mu \\
\text{s.t.} V^\pi(\mu) = E_s \sum_{t=1}^H \sum_{i=1}^{n_i} c_i | \mu \leq \hat{c}
\]

**Cost-model-irrelevant Abstraction**

\[ M_\phi = (S, A, \hat{P}, \hat{R}, \hat{\mu}, \hat{C}, \hat{\hat{c}}) \]

\[ \phi : S \rightarrow \hat{S} \]

\[ M = (S, A, \hat{P}, R, \mu, C, \hat{c}) \]

\[ \phi \text{ preserves the expected cost: } V^\pi_{\hat{c}}(\hat{\mu}) = V^\pi_C(\mu) \]

- The abstract policy \( \pi_\phi \) is **safe** but **might be suboptimal**.
- The ground policy \( \pi_0 \) **can reach optimality** but has no safety guarantees.

**Uncertainty set**

\[ \Sigma = \{ T \in P(S) | \| T(\cdot | s, a) - T(\cdot | s, a) \| \leq \sigma(s, a) \} \]

\( \Sigma \) contains the true transition function with high probability.

**Conservative policy**

Tight safety constraint until ground policy is safe in all probable CMDPs (\( \Sigma \)).

**Results**

Find more at: [https://tdsimao.github.io/publications/Simao2021alwayssafe/](https://tdsimao.github.io/publications/Simao2021alwayssafe/)