

Bridge

Michael Lu

February 4, 2021

Contents

1	Key Ideas	1
2	Notes	1
3	Notation	1

1 Key Ideas

- *failure states*

2 Notes

- introduce new Safety Bellman Equation
- Safety Q-learning algorithm converges to safe state-action value in finite MDP (what about conti. case?)
- with deterministic dynamics $\dot{x} = f(x, u)$, the dp associated with control problem can be the discrete-time Bellman equation

$$V(x) = \max_{u \in \mathcal{U}} r(x, u) + \gamma V(x + f(x, u)\Delta t)$$

3 Notation

- $\mathcal{K} :=$ set of states that are safe for all future times $t \geq 0$
- $V(x) := \supinf(l\Sigma_x)$