Planning Maintenance Activities on a Network of Bridges (An Overview)

Zachary Hamida Postdoc
James-A. Goulet Professeur

Polytechnique Montréal, Canada
Département des génies civil, géologique et des mines

April 30, 2021

Partenaire: MTQ

Outline

Context & Objectives

Planning Interventions

Reinforcement Learning for Planning

Challenges

Interventions & Infrastructures

Infrastructure Health:

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention:

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Types of Interventions:

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Types of Interventions:

- Preventative maintenance.

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Types of Interventions:

- Preventative maintenance.
- Routine maintenance.

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Types of Interventions:

- Preventative maintenance.
- Routine maintenance.
- Repairs.

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

Types of Interventions:

- Preventative maintenance.
- Routine maintenance.
- Repairs.
- Major activities (i.e., Replacement).

Source: MTQ, Manual of Inspection

Interventions & Infrastructures

Infrastructure Health: the capacity of the components to function and serve their roles.

Intervention: maintenance activity that helps sustain or improve the health state of infrastructures.

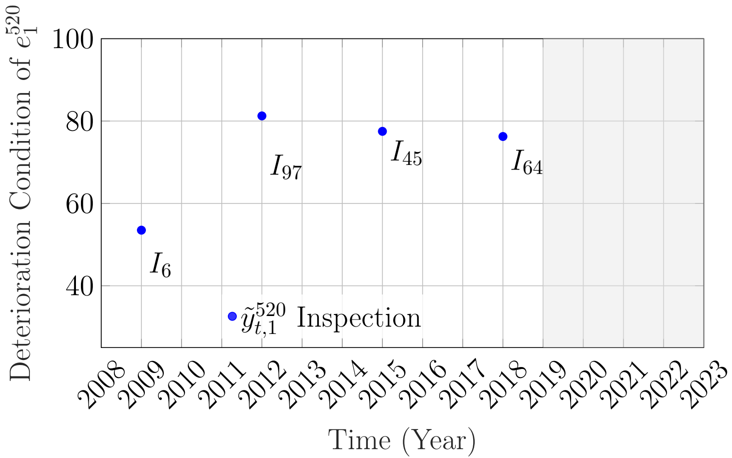
Types of Interventions:

- Preventative maintenance.
- Routine maintenance.
- Repairs.
- Major activities (i.e., Replacement).

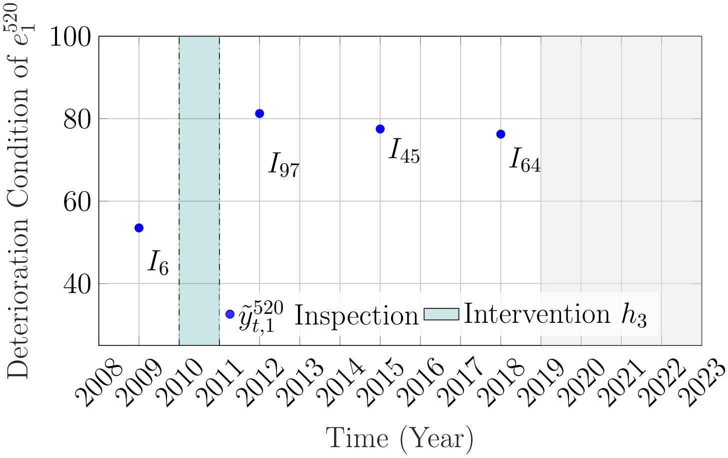
Preventative maintenance, routine maintenance and repairs: **Same element**

Source: MTQ, Manual of Inspection

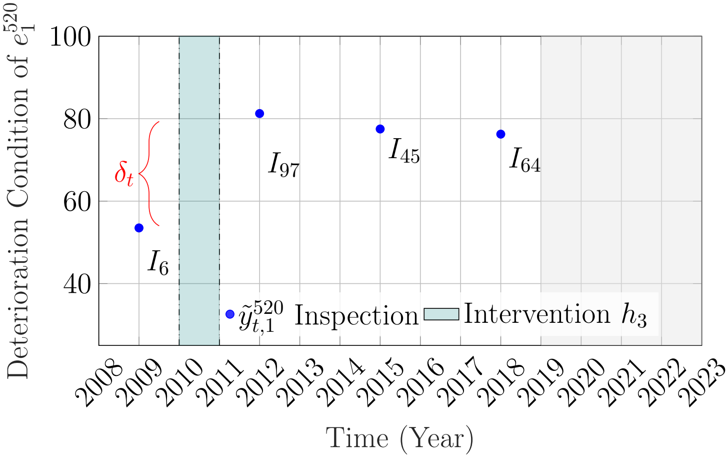
Example of an Intervention on Element - Repairs



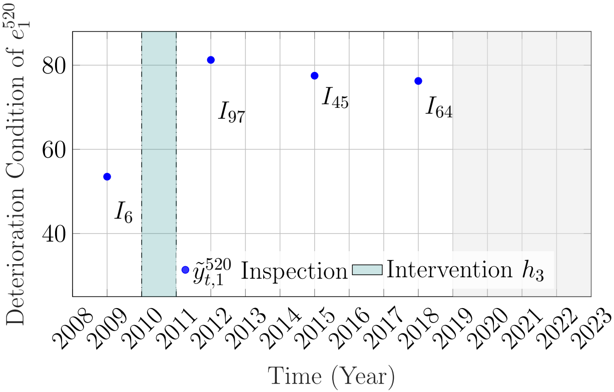
Example of an Intervention on Element - Repairs

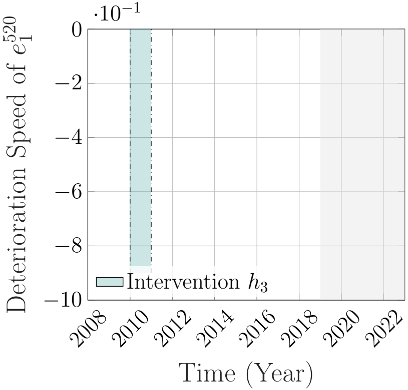


Example of an Intervention on Element - Repairs

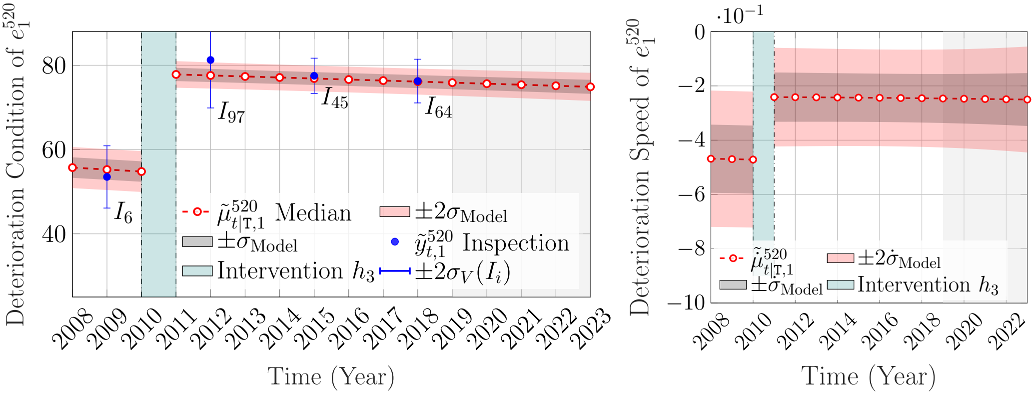


Example of an Intervention on Element - Repairs





Example of an Intervention on Element - Repairs



Objectives

- **Interventions planning** under budgetary constraints over time

Objectives

- **Interventions planning** under budgetary constraints over time
  - Maximize cost effectiveness.

Objectives

- **Interventions planning** under budgetary constraints over time
  - Maximize cost effectiveness.
  - Maximize the average health state of the network.

Cost Breakdown

Cost Breakdown

 $\longrightarrow$  Health state

Cost Breakdown

$\longrightarrow$

Health state

Maintenance Costs:

Cost Breakdown

$\longrightarrow$

Health state

Maintenance Costs:

- Intervention cost.

Cost Breakdown

 $\longrightarrow$  Health state

Maintenance Costs:

- Intervention cost.
  - Structural element type.

Cost Breakdown

$\longrightarrow$


Health state

Maintenance Costs:

- Intervention cost.
  - Structural element type.
  - Intervention category.

Cost Breakdown

$\rightarrow$

Health state

Maintenance Costs:

- Intervention cost.
  - Structural element type.
  - Intervention category.
  - Quantity of the element.

Cost Breakdown

$\longrightarrow$

Health state

Maintenance Costs:

- Intervention cost.
  - Structural element type.
  - Intervention category.
  - Quantity of the element.

- Maintenance of traffic and contingency costs.

Cost Breakdown

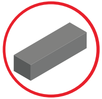
 $\longrightarrow$  Health state

Maintenance Costs:

- Intervention cost.
  - Structural element type.
  - Intervention category.
  - Quantity of the element.
- Maintenance of traffic and contingency costs.
  - From 0 to 3.5×10^6.

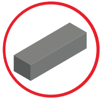
Decisions Hierarchy

Decisions Hierarchy

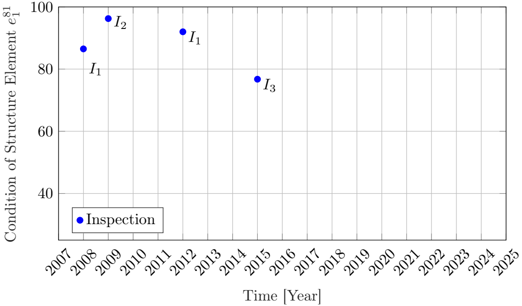


$$\begin{cases} \text{Time } \tau \\ \text{Type } h \end{cases}$$

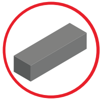
Decisions Hierarchy



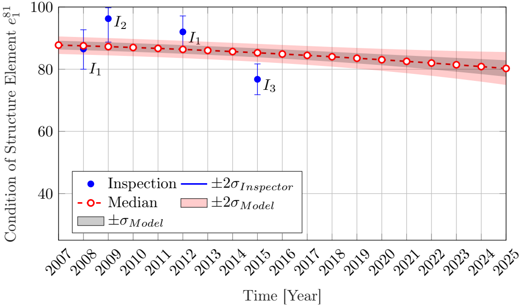
$$\begin{cases} \text{Time } \tau \\ \text{Type } h \end{cases}$$



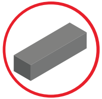
Decisions Hierarchy



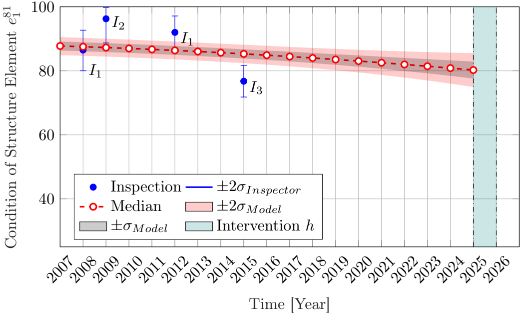
$$\begin{cases}\text{Time } \tau \\ \text{Type } h\end{cases}$$



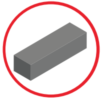
Decisions Hierarchy



$$\begin{cases} \text{Time } \tau \\ \text{Type } h \end{cases}$$



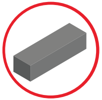
Decisions Hierarchy



$$\begin{cases} \text{Time } \tau \\ \text{Type } h \end{cases}$$

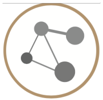
$$\begin{cases} \text{Time } \tau_b \\ \text{Number } \mathtt{E} \end{cases}$$

Decisions Hierarchy



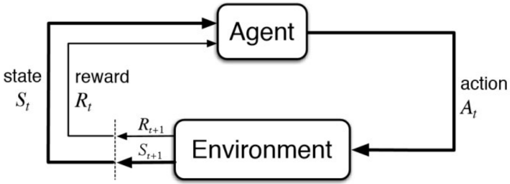
$\begin{cases} \text{Time } \tau \\ \text{Type } h \end{cases}$

$\begin{cases} \text{Time } \tau_b \\ \text{Number } \mathtt{E} \end{cases}$

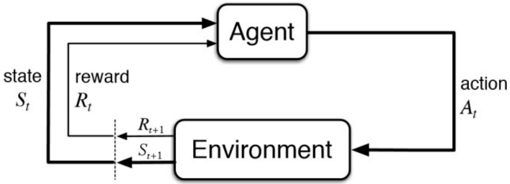


$\begin{cases} \$ \mid \text{Health State (i.e., sustain / improve)} \ \textbf{OR} \\ \text{B} \mid \$ \end{cases}$

RL Model

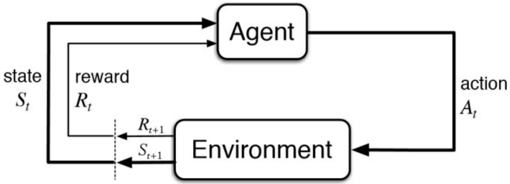


RL Model



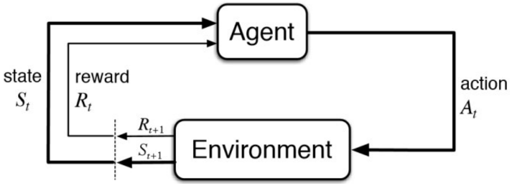
- ▷ Agent: decision maker.

RL Model



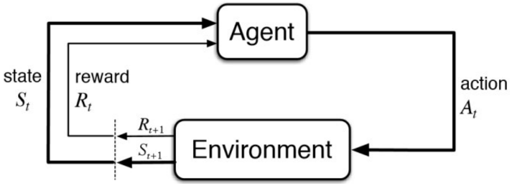
- ▷ Agent: decision maker.
- ▷ Environment: deterioration process.

RL Model



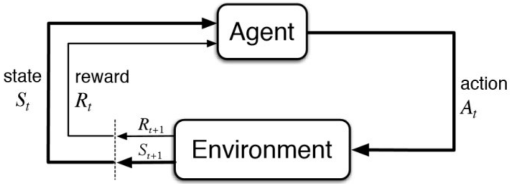
- ▷ Agent: decision maker.
- ▷ Environment: deterioration process.
- ▷ Actions: type of intervention.

RL Model



- ▷ Agent: decision maker.
- ▷ Environment: deterioration process.
- ▷ Actions: type of intervention.
- ▷ Rewards: interventions costs, budget, achieving the goal...etc.

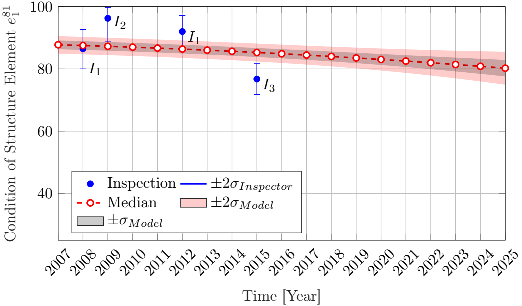
RL Model



- Agent: decision maker.
- Environment: deterioration process.
- Actions: type of intervention.
- Rewards: interventions costs, budget, achieving the goal...etc.

The goal of an agent is to maximize the total rewards in an episode

Element-Level Problem Formulation



Goal: extend the life span for e for $t = 120$ years $\equiv$ **Episode**.

Environment and Actions

- Environment: SSM/SSM-KR model.

Environment and Actions

- Environment: SSM/SSM-KR model.
  - State : $\{\mu_{t|T}, \dot{\mu}_{t|T}, \#$ previous interventions$\}$.
- Actions.

Environment and Actions

- Environment: SSM/SSM-KR model.
  - State : $\{\mu_{t|T}, \dot{\mu}_{t|T}, \#\text{ previous interventions}\}$.
- Actions.
  - 0: do nothing.

Environment and Actions

- Environment: SSM/SSM-KR model.
  - State : $\{\mu_{t|T}, \dot{\mu}_{t|T}, \#\text{ previous interventions}\}$.
- Actions.
  - 0: do nothing.
  - 1: preventive maintenance.

Environment and Actions

- Environment: SSM/SSM-KR model.
  - State : $\{\mu_{t|T}, \dot{\mu}_{t|T}, \#\text{ previous interventions}\}$.
- Actions.
  - 0: do nothing.
  - 1: preventive maintenance.
  - 2: routine maintenance.

Environment and Actions

- Environment: SSM/SSM-KR model.
  - State : $\{\mu_{t|T}, \dot{\mu}_{t|T}, \#\text{ previous interventions}\}$.
- Actions.
  - 0: do nothing.
  - 1: preventive maintenance.
  - 2: routine maintenance.
  - 3: repairs.

Reward Function

- Rewards & penalties.

Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.

Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.

Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.

Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.

Reward Function

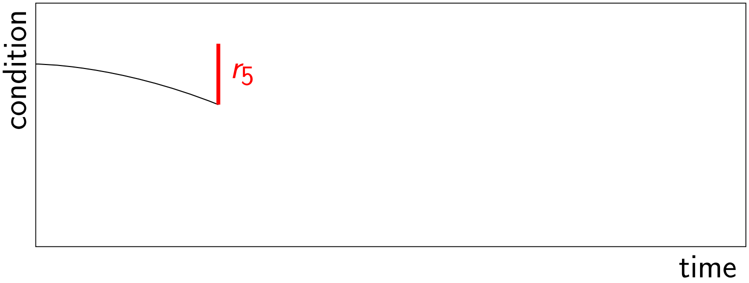
- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.

Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.

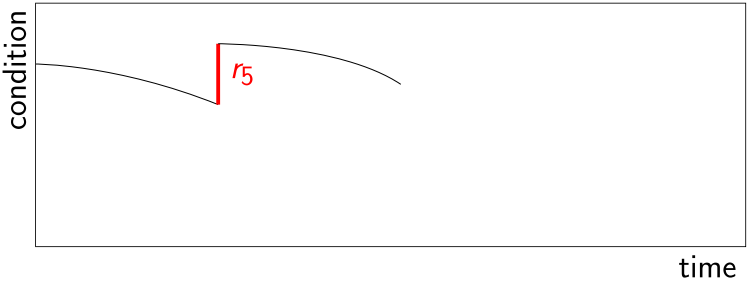
Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.



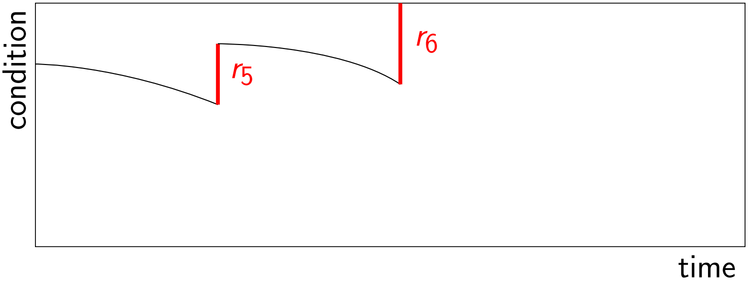
Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.



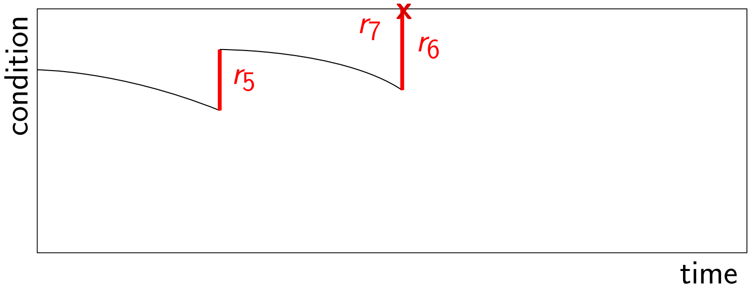
Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.



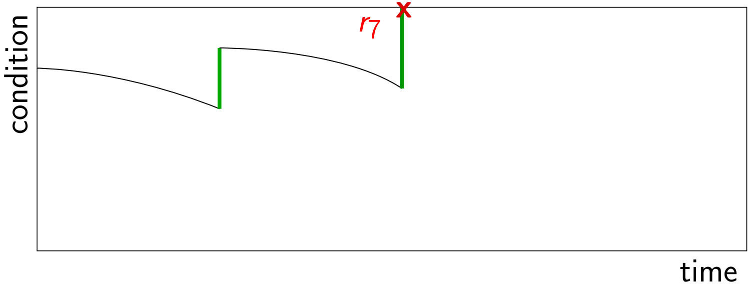
Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention costs $\rightarrow r_4, r_5, r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.



Reward Function

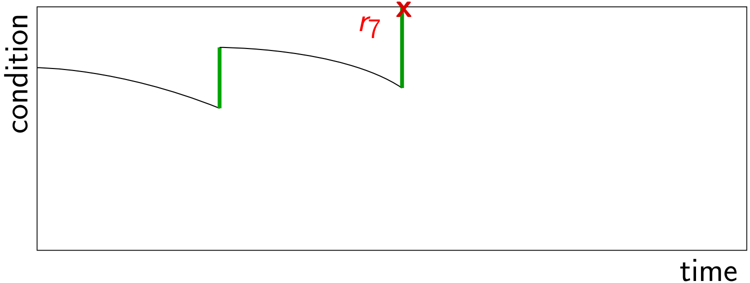
- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention effectiveness $\rightarrow \delta_1/r_4, \delta_2/r_5, \delta_3/r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.



Reward Function

- Rewards & penalties.
  - $\mu_{t(\text{terminal})} \in [75, 100] \rightarrow r_1$.
  - $\mu_{t(\text{terminal})} \in [50, 75] \rightarrow r_2$.
  - $\mu_{t(\text{terminal})} \in [25, 50] \rightarrow r_3$.
  - Intervention effectiveness $\rightarrow \delta_1/r_4, \delta_2/r_5, \delta_3/r_6$.
  - Bad interventions (early termination) $\rightarrow r_7$.

Convergence is either slow or unsatisfactory

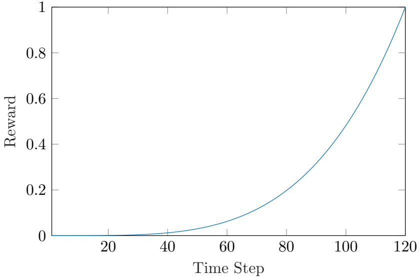


Reward Shaping

Reward Shaping: a technique to incorporate knowledge about the agent's goal.

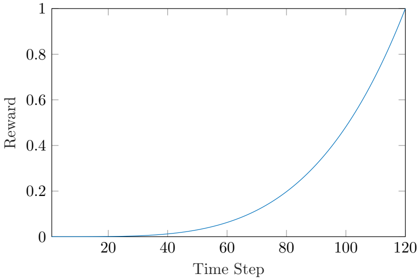
Reward Shaping

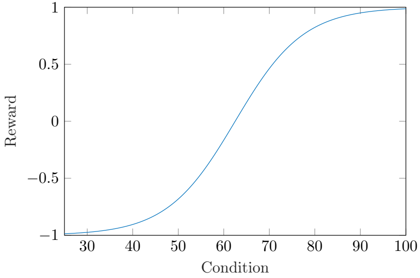
Reward Shaping: a technique to incorporate knowledge about the agent's goal.



Reward Shaping

Reward Shaping: a technique to incorporate knowledge about the agent's goal.





Q-Learning

The total rewards:

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots$$

The Q-value:

$$Q(s_t, a_t) = \mathbb{E}[R_t | s_t, a_t]$$

Q-Learning

The total rewards:

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots$$

The Q-value:

$$Q(s_t, a_t) = \mathbb{E}[R_t | s_t, a_t]$$

$$Q(s, a) = r(s, a, s') + \gamma \max_{a'} Q(s', a')$$

Q-Learning

The total rewards:

$$R_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \ldots$$

The Q-value:

$$Q(s_t, a_t) = \mathbb{E}[R_t | s_t, a_t]$$

$$Q(s, a) = Q(s, a) + \alpha[r(s, a, s') + \gamma \max_{a'} Q(s', a') - Q(s, a)]$$

Q-Learning

Algorithm 1 Q Learning for estimating π^*

1: Initialize $Q(s, a), \alpha, \gamma$

Q-Learning

Algorithm 2 Q Learning for estimating π^*

1: Initialize $Q(s,a), \alpha, \gamma$
2: **for** *episode* $= 1$ to N **do**

Q-Learning

Algorithm 3 Q Learning for estimating π^*

1: Initialize $Q(s, a), \alpha, \gamma$
2: **for** *episode* $= 1$ to N **do**
3: Initialize s

Q-Learning

Algorithm 4 Q Learning for estimating π^*

1: Initialize $Q(s, a), \alpha, \gamma$
2: **for** *episode* $= 1$ to N **do**
3: Initialize s
4: **for** step $= 1$ to terminal-state **do**

Q-Learning

Algorithm 5 Q Learning for estimating π^*

```
Initialize Q(s, a), α, γ
for episode = 1 to N do
    Initialize s
    for step = 1 to terminal-state do
        a ← select-action(Q, s, ε)
```

Q-Learning

Algorithm 6 Q Learning for estimating π^*

1: Initialize $Q(s, a), \alpha, \gamma$
2: **for** *episode* $= 1$ to N **do**
3: Initialize s
4: **for** step $= 1$ to terminal-state **do**
5: $a \leftarrow$ select-action(Q, s, ϵ)
6: $s', r \leftarrow$ agent-environment(a)

Q-Learning

Algorithm 7 Q Learning for estimating π^*

Initialize $Q(s,a), \alpha, \gamma$
for *episode* $= 1$ to N **do**
  Initialize s
  for step $= 1$ to terminal-state **do**
    $a \leftarrow$ select-action(Q, s, ϵ)
    $s', r \leftarrow$ agent-environment(a)
    $Q(s,a) = Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$

Q-Learning

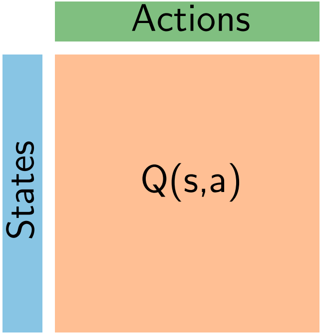
Algorithm 8 Q Learning for estimating π^*

1: Initialize $Q(s, a), \alpha, \gamma$
2: **for** *episode* $= 1$ to N **do**
3: Initialize s
4: **for** step $= 1$ to terminal-state **do**
5: $a \leftarrow$ select-action(Q, s, ϵ)
6: $s', r \leftarrow$ agent-environment(a)
7: $Q(s, a) = Q(s, a) + \alpha \left[r + \gamma \max_{a'} Q(s', a') - Q(s, a) \right]$
8: $s \leftarrow s'$

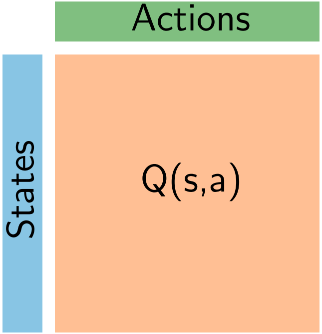
Q-Learning

Algorithm 9 Q Learning for estimating π^*

Initialize $Q(s,a), \alpha, \gamma$
for *episode* $= 1$ to N **do**
    Initialize s
    for step $= 1$ to terminal-state **do**
        $a \leftarrow \text{select-action}(Q, s, \epsilon)$
        $s', r \leftarrow \text{agent-environment}(a)$
        $Q(s,a) = Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$
        $s \leftarrow s'$
    end for
end for

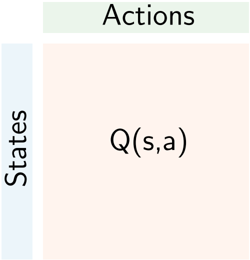


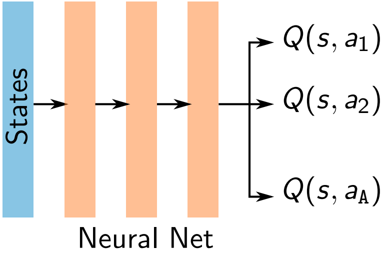
$$Q(s,a) = Q(s,a) + \alpha\left[r + \gamma \max_{a'} Q(s',a') - Q(s,a)\right]$$

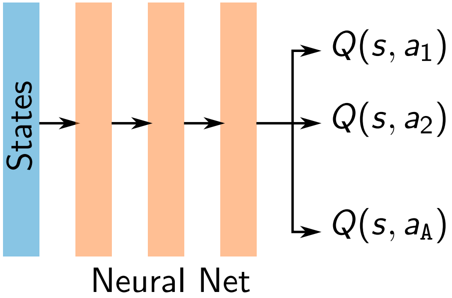


$$Q(s,a) = Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$$

High computational demand.


Actions
States
Q(s,a)


States
$Q(s, a_1)$
$Q(s, a_2)$
$Q(s, a_{\mathtt{A}})$
Neural Net



$$\mathcal{L} = \| \underbrace{r + \gamma \max_{a'} Q(s', a')}_{\text{Target}} - \underbrace{Q(s, a)}_{\text{Predicted}} \|^2$$

Q-Learning

Algorithm 10 Deep Q Learning with Experience Replay for estimating π^*

1: Initialize Memory Size, Batch Size, ϵ, γ

Q-Learning

Algorithm 11 Deep Q Learning with Experience Replay for estimating π^*

1: Initialize Memory Size, Batch Size, ϵ, γ
2: **for** *episode* $= 1$ to N **do**

Q-Learning

Algorithm 12 Deep Q Learning with Experience Replay for estimating π^*

1: Initialize Memory Size, Batch Size, ϵ, γ
2: **for** *episode* $= 1$ to N **do**
3: Initialize s

Q-Learning

Algorithm 13 Deep Q Learning with Experience Replay for estimating π^*

1: Initialize Memory Size, Batch Size, ϵ, γ
2: **for** *episode* $= 1$ to N **do**
3: Initialize s
4: **for** step $= 1$ to terminal-state **do**

Q-Learning

Algorithm 14 Deep Q Learning with Experience Replay for estimating π^*

1: Initialize Memory Size, Batch Size, ϵ, γ
2: **for** *episode* $= 1$ to N **do**
3: Initialize s
4: **for** step $= 1$ to terminal-state **do**
5: $a \leftarrow$ select-action(Q, s, ϵ)

Q-Learning

Algorithm 15 Deep Q Learning with Experience Replay for estimating π^*

Initialize Memory Size, Batch Size, ϵ, γ
for *episode* $= 1$ to N **do**
  Initialize s
  for step $= 1$ to terminal-state **do**
    $a \leftarrow$ select-action(Q, s, ϵ)
    $s', r,$ terminal $\leftarrow$ agent-environment(a)

Q-Learning

Algorithm 16 Deep Q Learning with Experience Replay for estimating π^*

Initialize Memory Size, Batch Size, ϵ, γ
for $episode = 1$ to N **do**
    Initialize s
    for step $= 1$ to terminal-state **do**
        $a \leftarrow$ select-action(Q, s, ϵ)
        $s', r,$ terminal $\leftarrow$ agent-environment(a)
        Memory $\leftarrow$ s, a, r, s', terminal

Q-Learning

Algorithm 17 Deep Q Learning with Experience Replay for estimating π^*

Initialize Memory Size, Batch Size, ϵ, γ
for *episode* $= 1$ to N **do**
  Initialize s
  for step $= 1$ to terminal-state **do**
    $a \leftarrow$ select-action(Q, s, ϵ)
    $s', r,$ terminal $\leftarrow$ agent-environment(a)
    Memory $\leftarrow$ s, a, r, s', terminal
    Sample n experience from Memory

Q-Learning

Algorithm 18 Deep Q Learning with Experience Replay for estimating π^*

```
Initialize Memory Size, Batch Size, ε, γ
for episode = 1 to N do
    Initialize s
    for step = 1 to terminal-state do
        a ← select-action(Q, s, ε)
        s', r, terminal ← agent-environment(a)
        Memory ← s, a, r, s', terminal
        Sample n experience from Memory
        y = { r,  if terminal,
            { r + γ max_{a'} Q(s', a'; θ')
        Minimize ||y − Q(s, a; θ)||^2
```

$$y = \begin{cases} r, & \text{if terminal,} \\ r + \gamma \max_{a'} Q(s', a'; \theta') \end{cases}$$

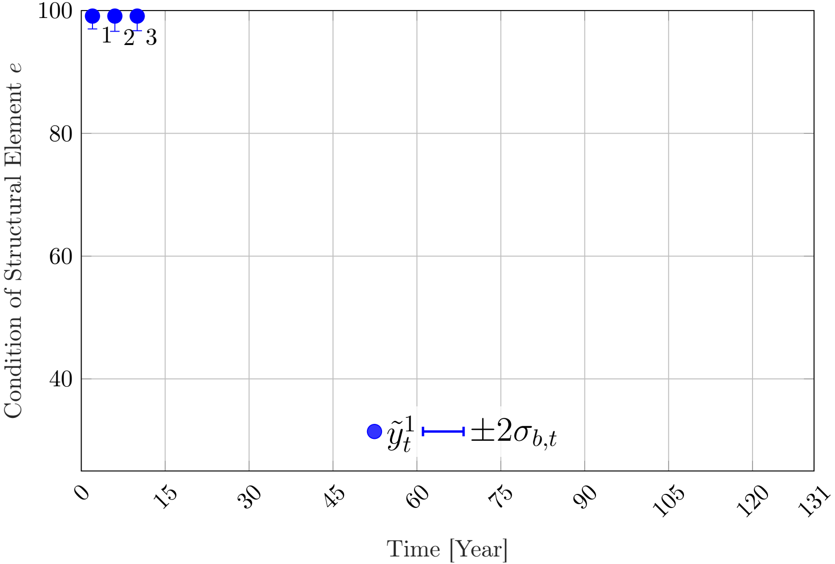
$$\text{Minimize } \|y - Q(s, a; \theta)\|^2$$

Q-Learning

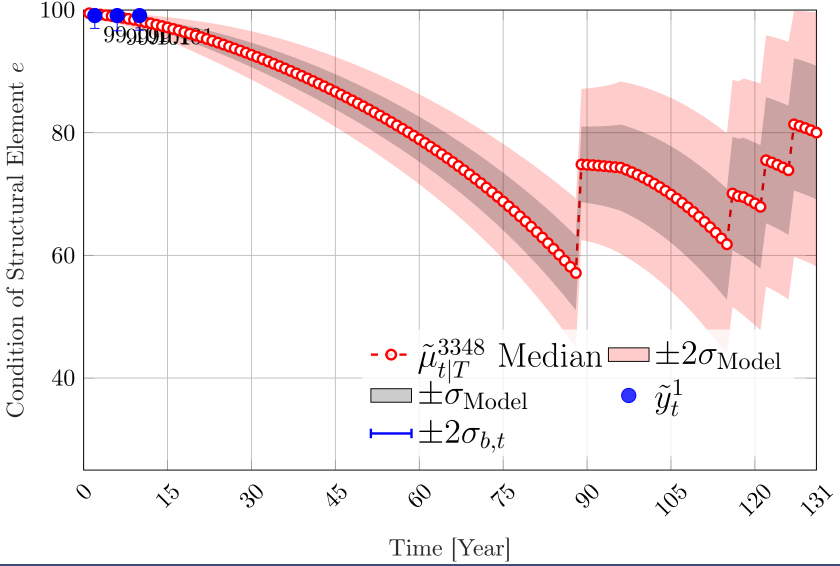
Algorithm 19 Deep Q Learning with Experience Replay for estimating π^*

Initialize Memory Size, Batch Size, ϵ, γ
for *episode* $= 1$ to N **do**
  Initialize s
  for step $= 1$ to terminal-state **do**
    $a \leftarrow$ select-action(Q, s, ϵ)
    $s', r,$ terminal $\leftarrow$ agent-environment(a)
    Memory $\leftarrow\; s, a, r, s'$, terminal
    Sample n experience from Memory
    $y = \begin{cases} r, & \text{if terminal,} \\ r + \gamma \max_{a'} Q(s', a'; \theta') \end{cases}$
    Minimize $\|y - Q(s, a; \theta)\|^2$
  end for
end for

Preliminary Results



Preliminary Results



Challenges:

- Convergence to sub-optimal / unrealstic solutions.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.
- Type of intervention is not determined.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.
- Type of intervention is not determined.
  - Intervention costs vary significantly within each intervention category.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.
- Type of intervention is not determined.
  - Intervention costs vary significantly within each intervention category.
- Determine the total long-term budget per structural element / bridge.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.
- Type of intervention is not determined.
  - Intervention costs vary significantly within each intervention category.
- Determine the total long-term budget per structural element / bridge.
- Online deterioration analyses in the environment.

Challenges:

- Convergence to sub-optimal / unrealstic solutions.
- Account for need-based interventions.
- Type of intervention is not determined.
  - Intervention costs vary significantly within each intervention category.
- Determine the total long-term budget per structural element / bridge.
- Online deterioration analyses in the environment.
- Network-level analyses.