Planning Interventions for Infrastructures Using Hierarchical Reinforcement Learning

Zachary Hamida Postdoc
James-A. Goulet Professeur

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

July 29, 2025

Funding:
Transportation Ministry of Quebec (MTQ)

Outline

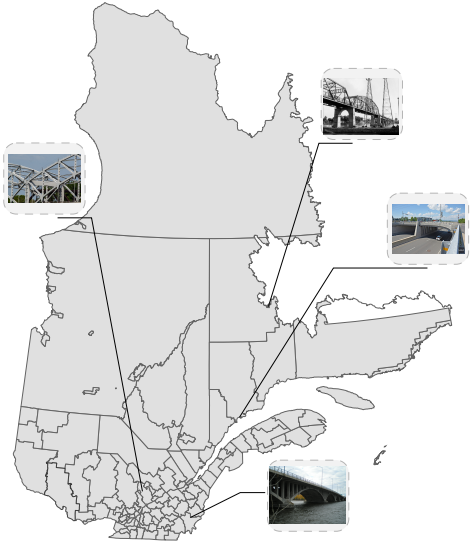
Recap & Objectives

Maintenance Framework

Element-level (Verification)

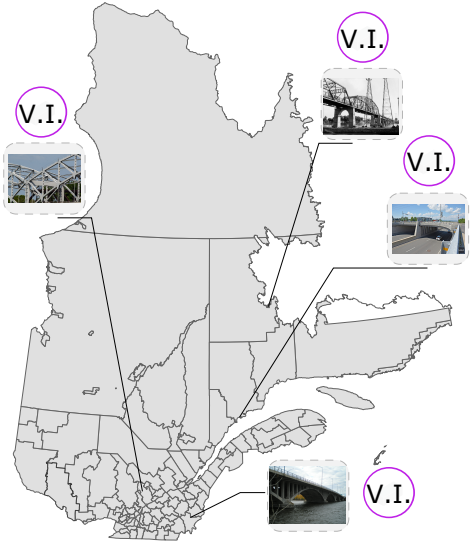
Structure & Network Planning

Quick Recap



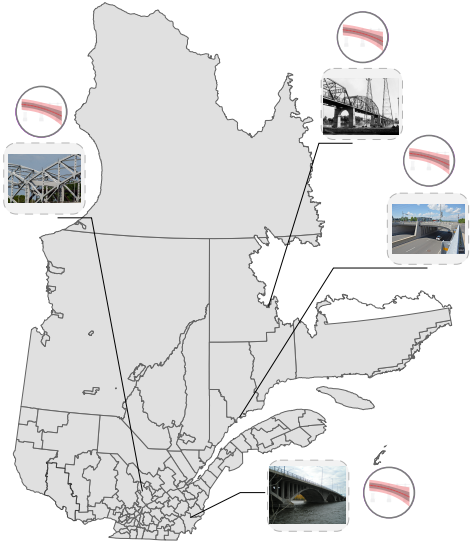
Quick Recap

- ▷ Element-level monitoring



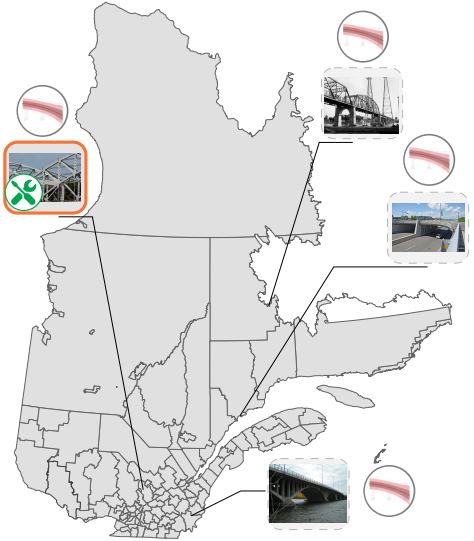
Quick Recap

- ▷ Element-level monitoring
- ▷ Deterioration model:
  - ▷ Element → Bridge.



Quick Recap

- ▷ Element-level monitoring
- ▷ Deterioration model:
  - ▷ Element → Bridge.



Quick Recap

- ▷ Element-level monitoring
- ▷ Deterioration model:
  - ▷ Element → Bridge.

Research Objective

Research Objective

- Design an approach for modelling network-scale decision-making.

Research Objective

- ▷ Design an approach for modelling network-scale decision-making.
- ▷ Develop a solver for network-scale decision-making:

Research Objective

- ▷ Design an approach for modelling network-scale decision-making.
- ▷ Develop a solver for network-scale decision-making:
  - ▷ Define the cost structure for maintenance actions.

Research Objective

- ▷ Design an approach for modelling network-scale decision-making.
- ▷ Develop a solver for network-scale decision-making:
  - ▷ Define the cost structure for maintenance actions.
  - ▷ Identify a long-term cost-effective maintenance plan.

Research Objective

- ▷ Design an approach for modelling network-scale decision-making.
- ▷ Develop a solver for network-scale decision-making:
  - ▷ Define the cost structure for maintenance actions.
  - ▷ Identify a long-term cost-effective maintenance plan.
- ▷ Validate the performance of the new decision-making systems.

Network-Scale Planning Challenges

Network-Scale Planning Challenges

Scale :

Network-Scale Planning Challenges

Scale :

$\underbrace{\text{Elements}}_{\times 10^5}$

Network-Scale Planning Challenges

Scale :  $\times$ 

Elements $\times 10^5$ Actions

Network-Scale Planning Challenges

Scale :  $\times$ 

Elements $\times 10^5$ Actions

Stochasticity :

Network-Scale Planning Challenges

Scale :  $\times$ 

Elements $\times 10^5$ Actions

Stochasticity : 

Deterioration Model

Network-Scale Planning Challenges

Scale :  $\times$ 

Elements $\times 10^5$ Actions

Stochasticity :  $+$ 

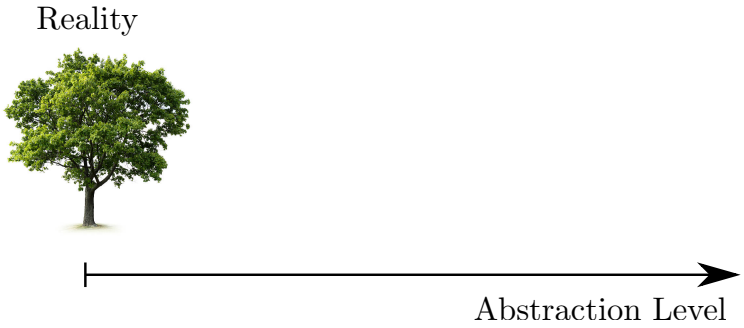
Deterioration Model Effect of Interventions

State Abstraction

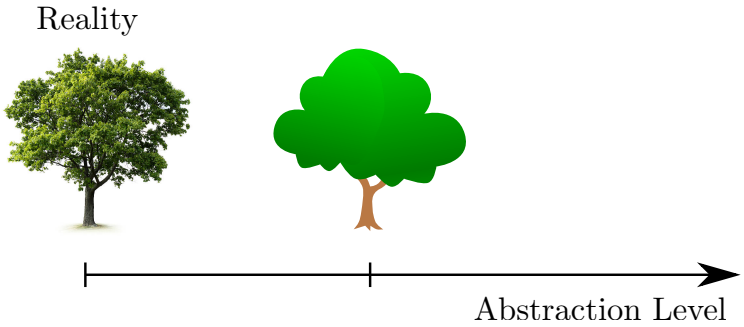
State Abstraction

Reality

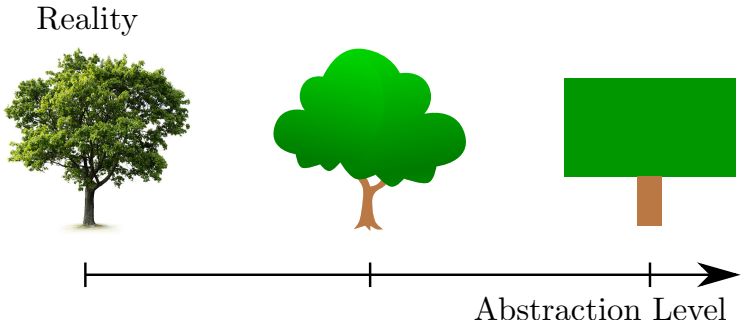
State Abstraction



State Abstraction



State Abstraction



Temporal Abstraction

Temporal Abstraction

$\mathcal{A}$: $\{a_0$: stay at home, a_1 : go to work $\}$

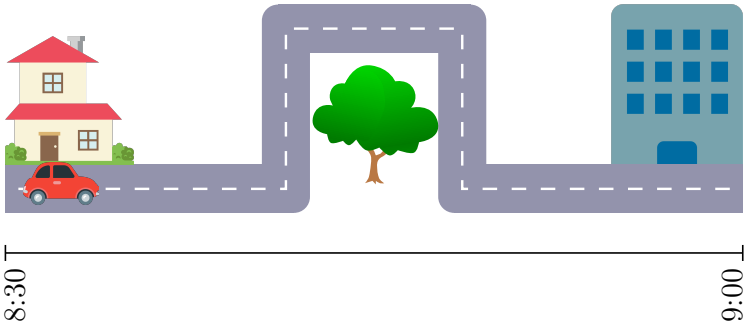
Temporal Abstraction

$\mathcal{A}$: $\{a_0$: stay at home, a_1 : **go to work**$\}$

Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \; a_1 : \textbf{go to work}\}$

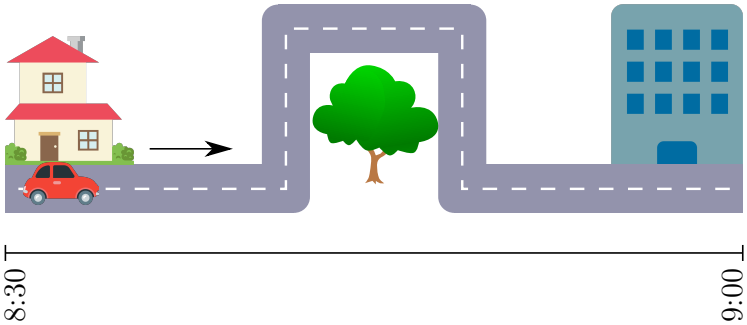
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \quad a_1 : \textbf{go to work}\}$

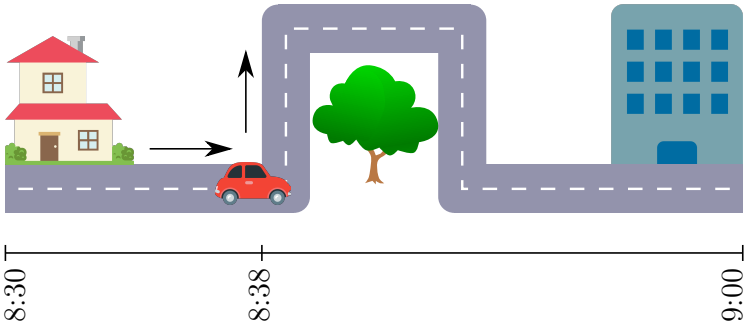
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \; a_1 : \textbf{go to work}\}$

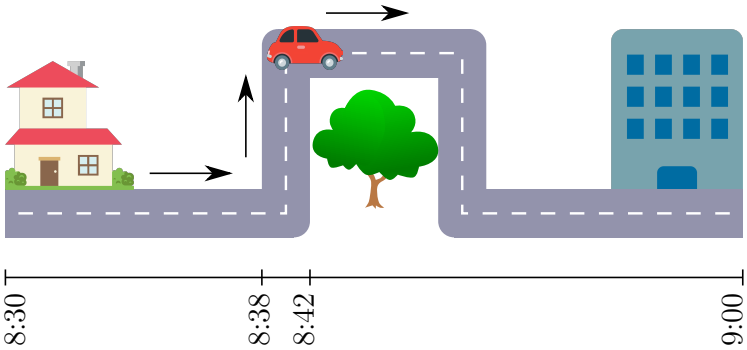
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \; a_1 : \textbf{go to work}\}$

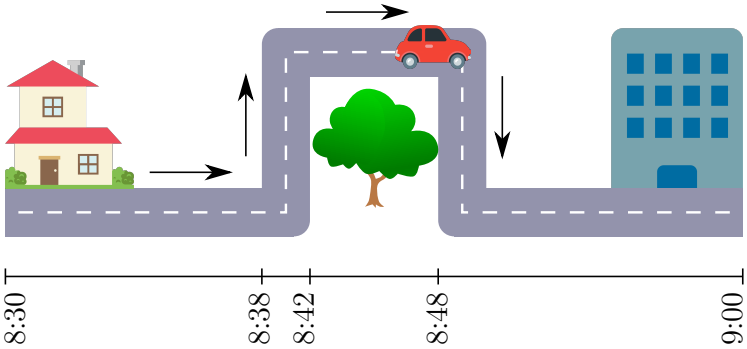
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \quad a_1 : \textbf{go to work}\}$

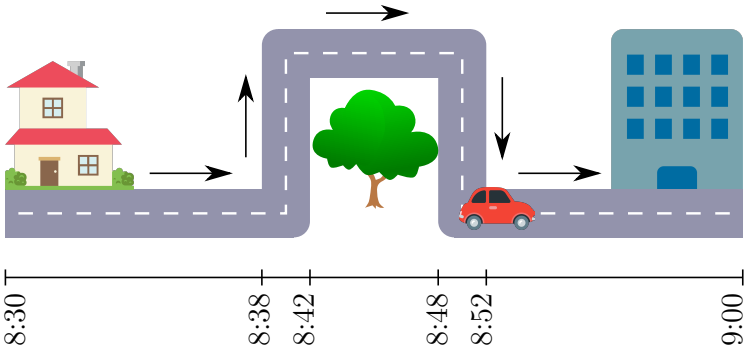
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

$\mathcal{A} : \{a_0 : \text{stay at home}, \quad a_1 : \textbf{go to work}\}$

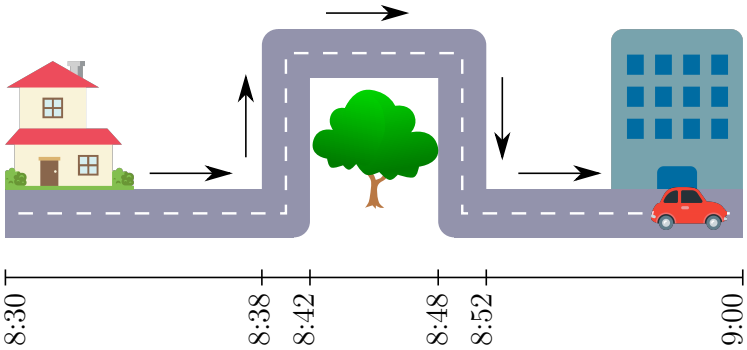
$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$



Temporal Abstraction

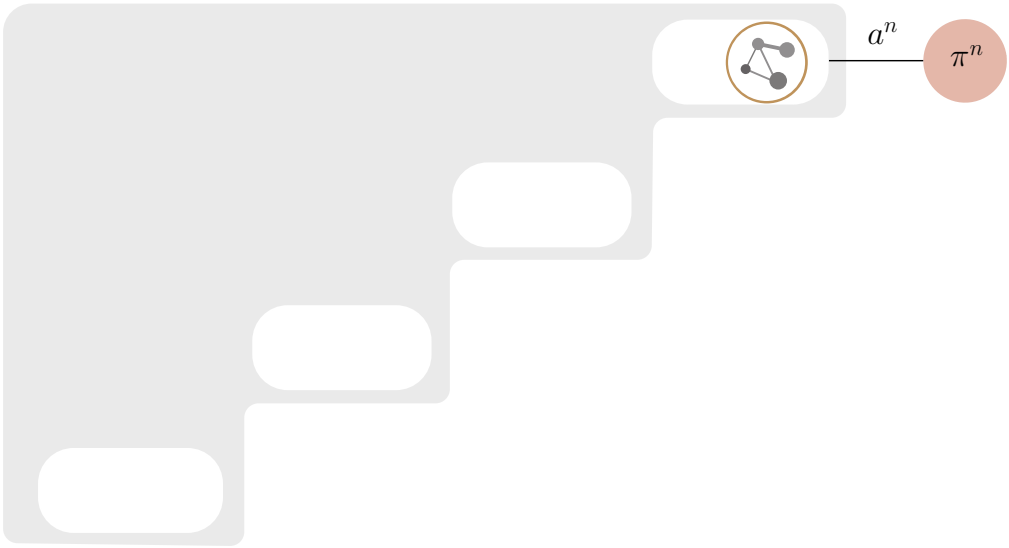
$\mathcal{A}$: $\{a_0$: stay at home, a_1 : **go to work**$\}$

$\hookrightarrow \mathcal{A}^{a_1} : \{\uparrow, \downarrow, \rightarrow, \leftarrow\}$

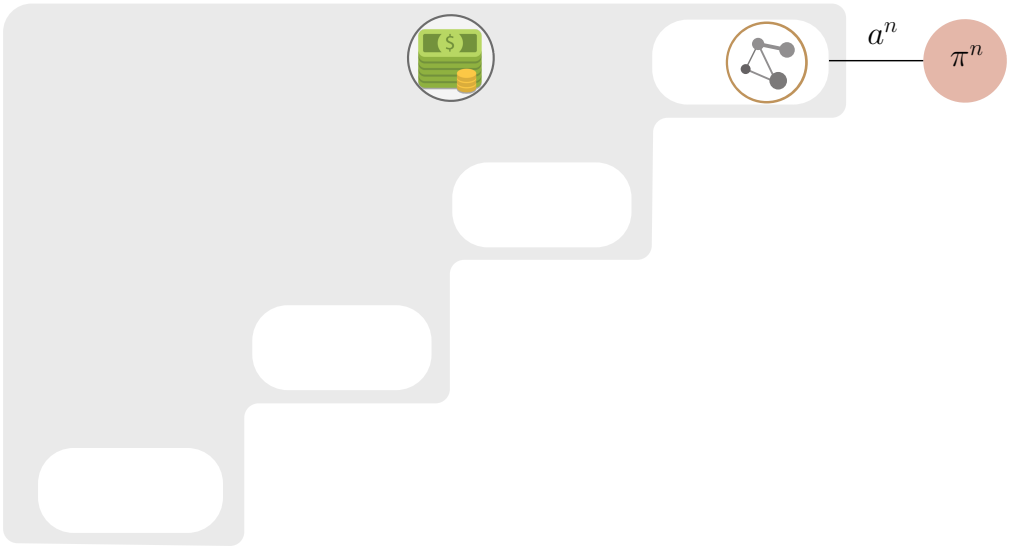


Network-Scale Maintenance Framework

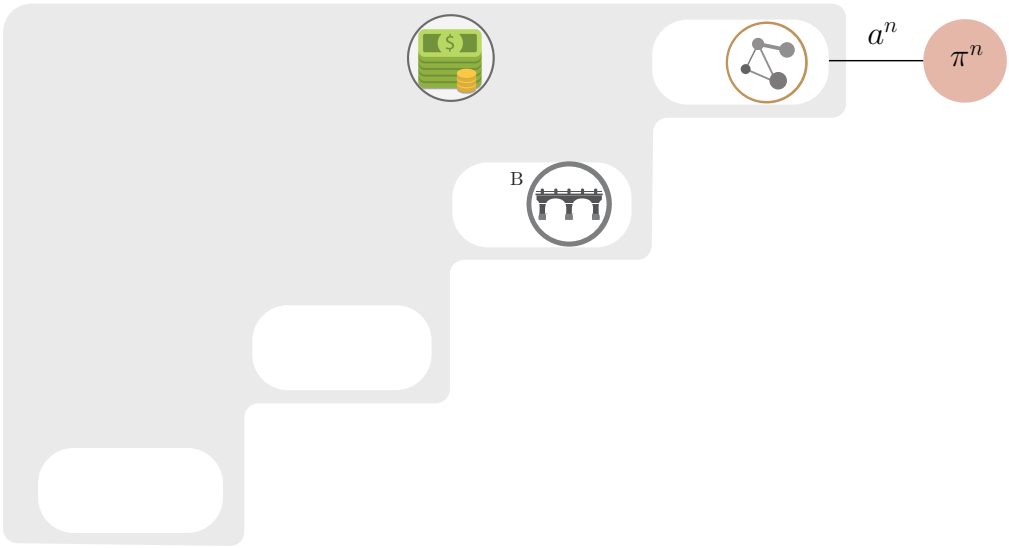
Network-Scale Maintenance Framework



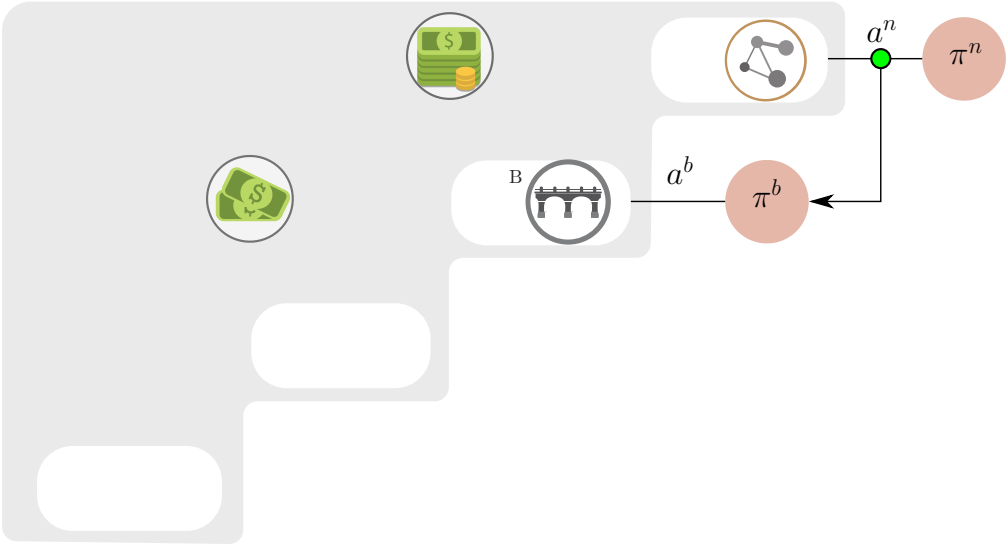
Network-Scale Maintenance Framework



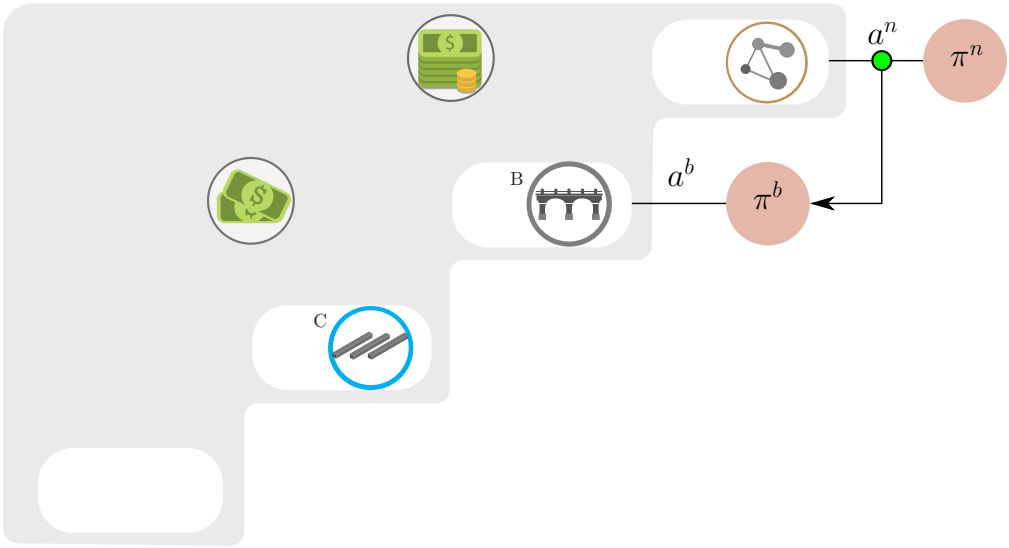
Network-Scale Maintenance Framework



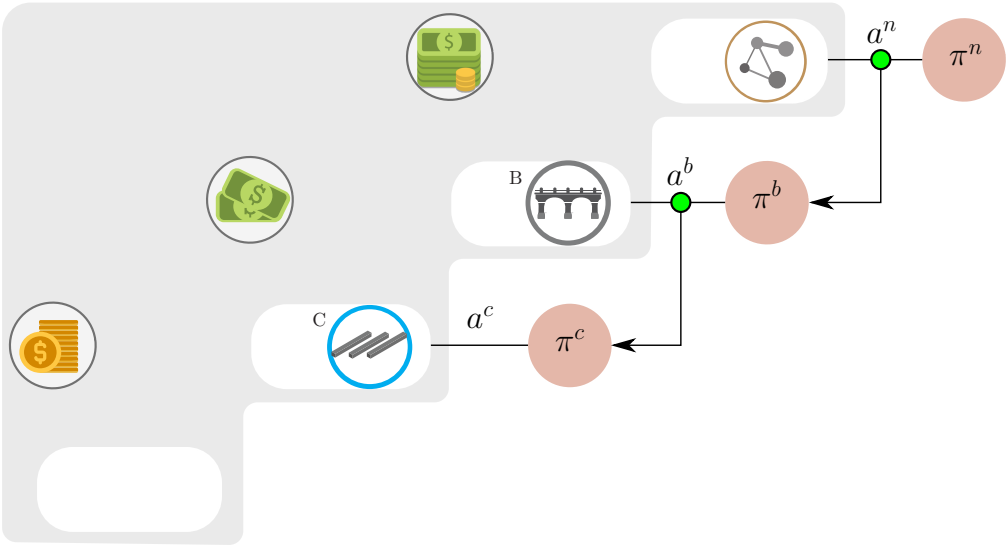
Network-Scale Maintenance Framework



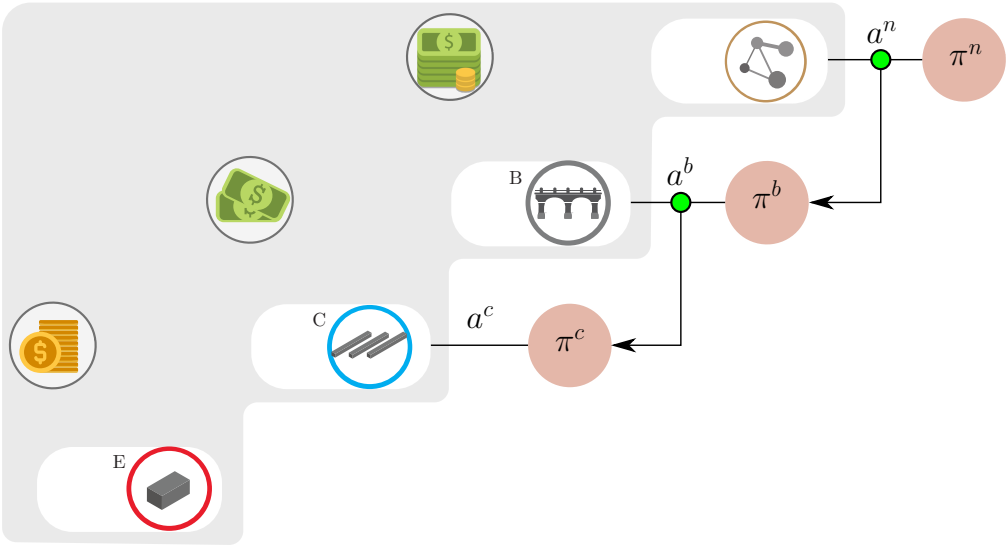
Network-Scale Maintenance Framework



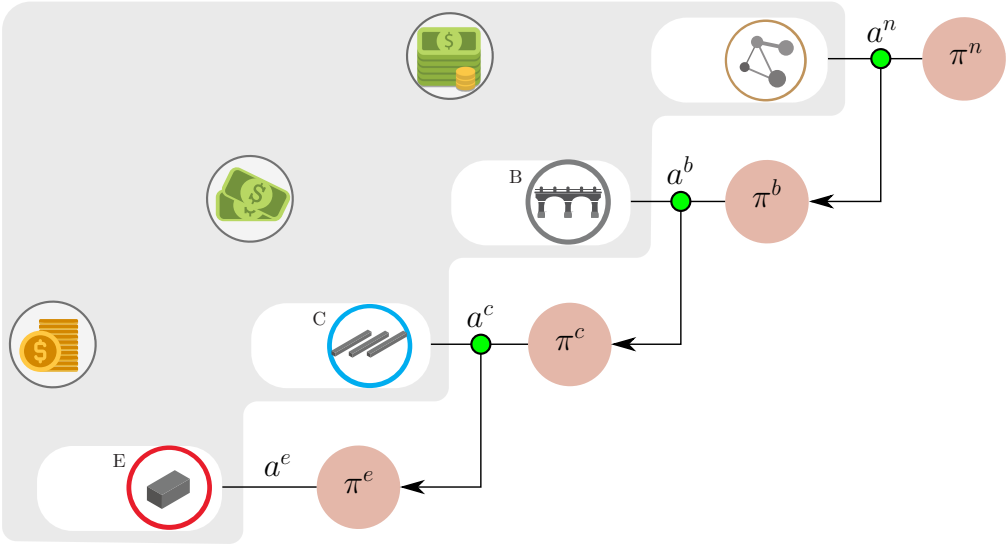
Network-Scale Maintenance Framework



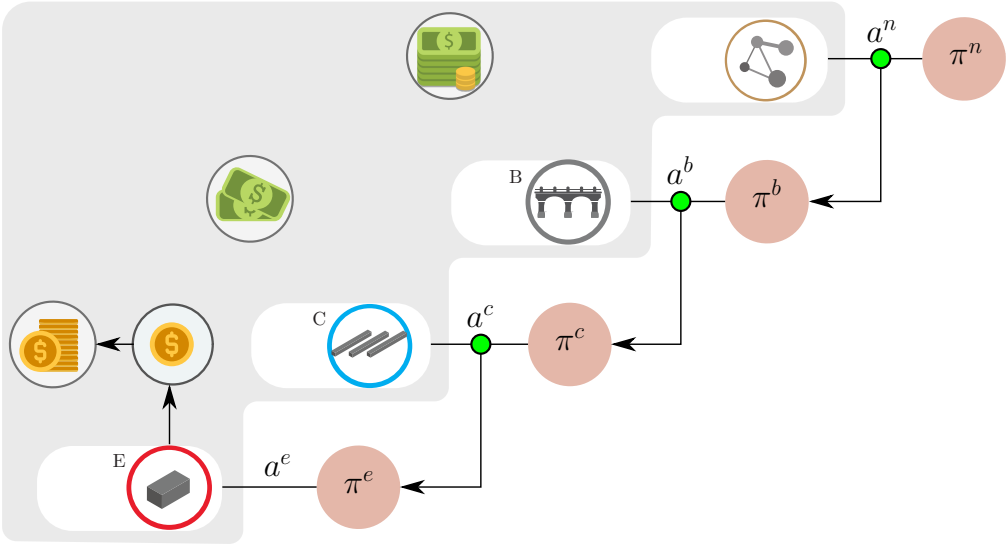
Network-Scale Maintenance Framework



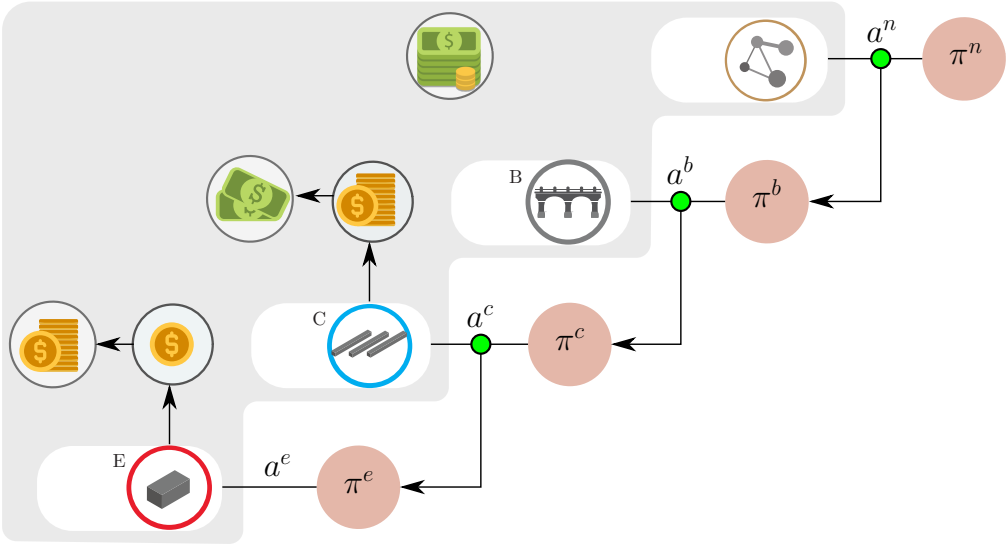
Network-Scale Maintenance Framework



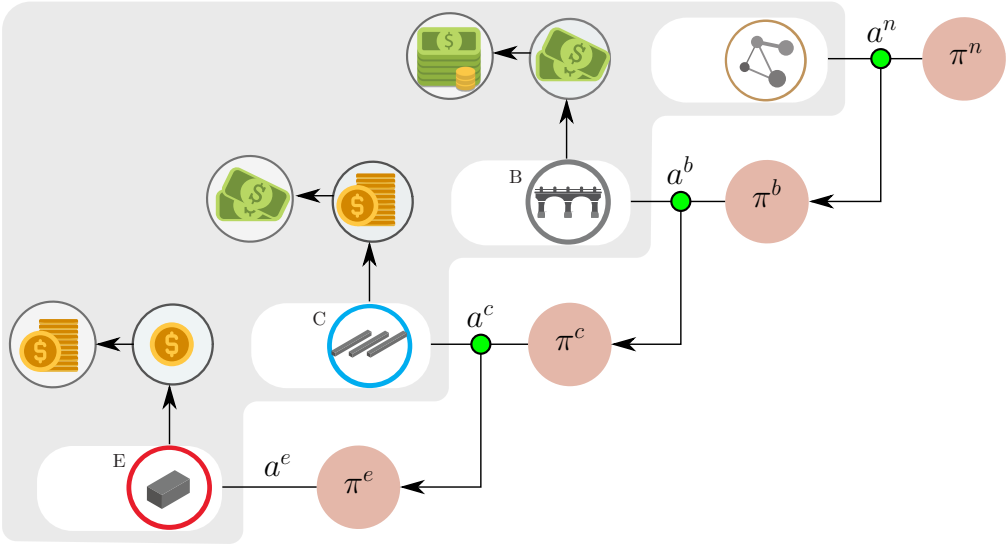
Network-Scale Maintenance Framework



Network-Scale Maintenance Framework



Network-Scale Maintenance Framework



State Representation

State Representation

Network State :

State Representation

Network State : $\left\{ \underbrace{\text{  }}_{\text{Network}} \right\}$

State Representation

$$\text{Network State} : \left\{ \underbrace{\text{  }}_{\text{Bridge \#001}}, \underbrace{\text{  }}_{\text{Bridge \#002}}, \ldots, \underbrace{\text{  }}_{\text{Bridge \#1}K} \right\}$$

State Representation

Network State : $\left\{ \underbrace{\text{  }}_{\text{Bridge \#001}}, \underbrace{\text{  }}_{\text{Bridge \#002}}, \ldots, \underbrace{\text{  }}_{\text{Bridge \#1}K} \right\}$

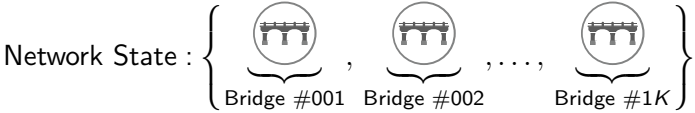
Bridge State :

State Representation

Network State : $\left\{ \underbrace{\text{  }}_{\text{Bridge \#001}}, \underbrace{\text{  }}_{\text{Bridge \#002}}, \ldots, \underbrace{\text{  }}_{\text{Bridge \#1}K} \right\}$

Bridge State : $\left\{ \underbrace{\text{  }}_{\text{Bridge}} \right\}$

State Representation



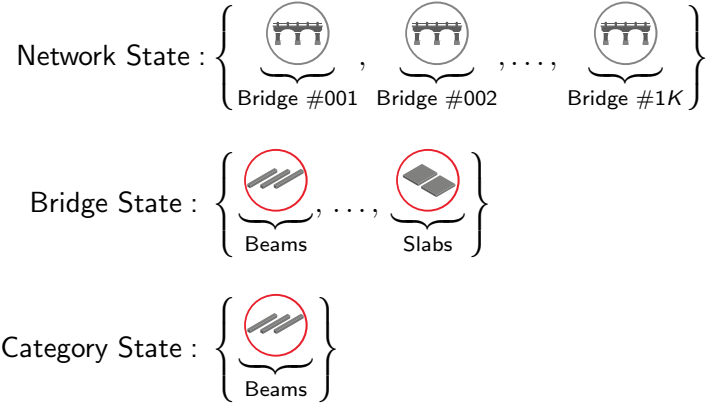
State Representation

Network State : $\left\{ \underbrace{\text{  }}_{\text{Bridge \#001}}, \underbrace{\text{  }}_{\text{Bridge \#002}}, \ldots, \underbrace{\text{  }}_{\text{Bridge \#1}K} \right\}$

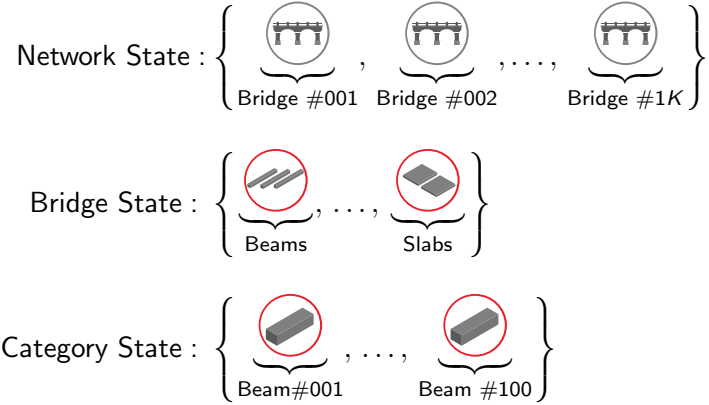
Bridge State : $\left\{ \underbrace{\text{  }}_{\text{Beams}}, \ldots, \underbrace{\text{  }}_{\text{Slabs}} \right\}$

Category State :

State Representation



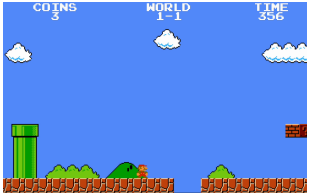
State Representation



Reinforcement Learning (RL)

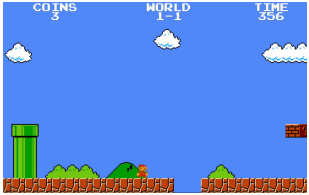
Source: Google Images

Reinforcement Learning (RL)



Source: Google Images

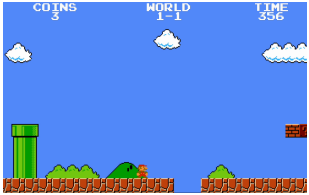
Reinforcement Learning (RL)

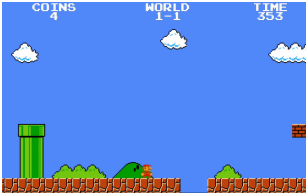


Bad decision (-1)

Source: Google Images

Reinforcement Learning (RL)

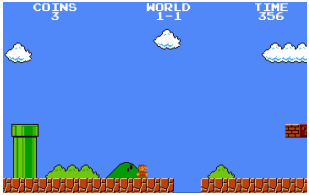




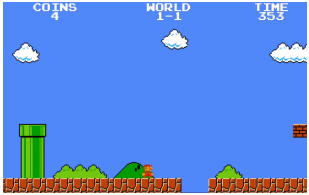
Bad decision (-1)

Source: Google Images

Reinforcement Learning (RL)



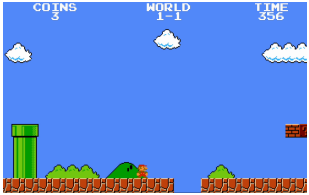
Bad decision (-1)



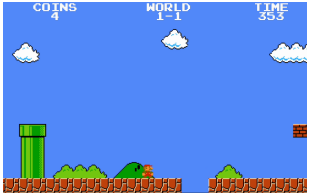
Good decision (+1)

Source: Google Images

Reinforcement Learning (RL)



Bad decision (-1)

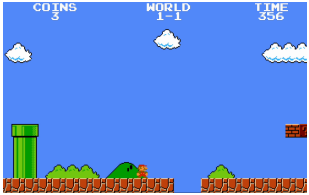


Good decision (+1)

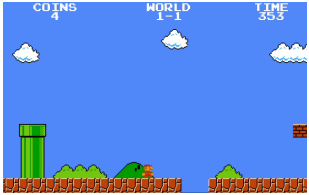
▷ Agent: Mario.

Source: Google Images

Reinforcement Learning (RL)



Bad decision (-1)

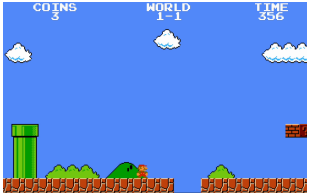


Good decision (+1)

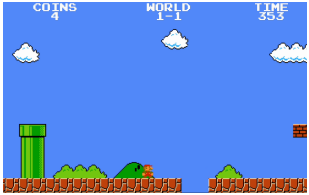
- ▷ Agent: Mario.
- ▷ Environnement: Mario World.

Source: Google Images

Reinforcement Learning (RL)



Bad decision (-1)

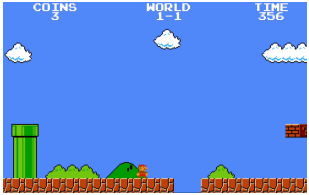


Good decision (+1)

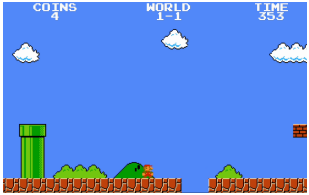
- ▷ Agent: Mario.
- ▷ Environnement: Mario World.
- ▷ Actions: Movement.

Source: Google Images

Reinforcement Learning (RL)



Bad decision (-1)

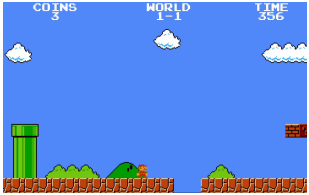


Good decision (+1)

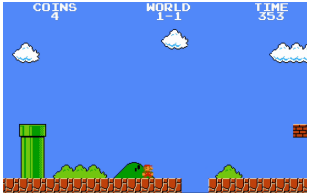
- ▷ Agent: Mario.
- ▷ Environnement: Mario World.
- ▷ Actions: Movement.
- ▷ Rewards: Win the game.

Source: Google Images

Reinforcement Learning (RL)



Bad decision (-1)



Good decision (+1)

- ▷ Agent: Mario.
- ▷ Environnement: Mario World.
- ▷ Actions: Movement.
- ▷ Rewards: Win the game.

RL Goal: learn a policy that maximizes the total expected discounted rewards.

Source: Google Images

Toy-Problem

Toy-Problem

▷ Problem: Find an optimal threshold to trigger an intervention.

Toy-Problem

▷ Problem: Find an optimal threshold to trigger an intervention.
▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.

Toy-Problem

- ▷ Problem: Find an optimal threshold to trigger an intervention.
- ▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.
- ▷ Actions: $\mathcal{A}^e = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Toy-Problem

- ▷ Problem: Find an optimal threshold to trigger an intervention.
- ▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.
- ▷ Actions: $\mathcal{A}^e = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Grid Search vs. Reinforcement Learning

Toy-Problem

▷ Problem: Find an optimal threshold to trigger an intervention.
▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.
▷ Actions: $\mathcal{A}^e = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Grid Search vs. Reinforcement Learning

Grid Search Setup:

Toy-Problem

▷ Problem: Find an optimal threshold to trigger an intervention.
▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.
▷ Actions: $\mathcal{A}^e = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Grid Search vs. Reinforcement Learning

Grid Search Setup:

▷ Planning duration $T = 100$ years.

Toy-Problem

▷ Problem: Find an optimal threshold to trigger an intervention.
▷ Environment: Element-level $\boldsymbol{s}_t^e = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}]$.
▷ Actions: $\mathcal{A}^e = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

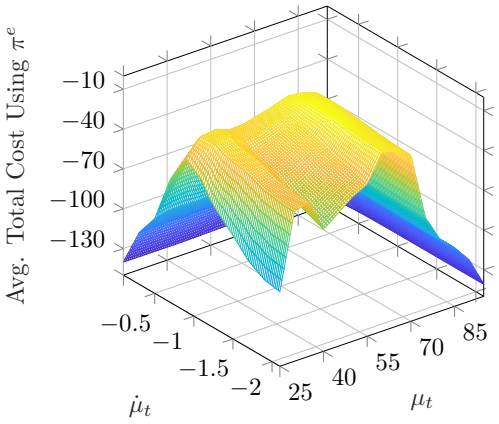
Grid Search vs. Reinforcement Learning

Grid Search Setup:

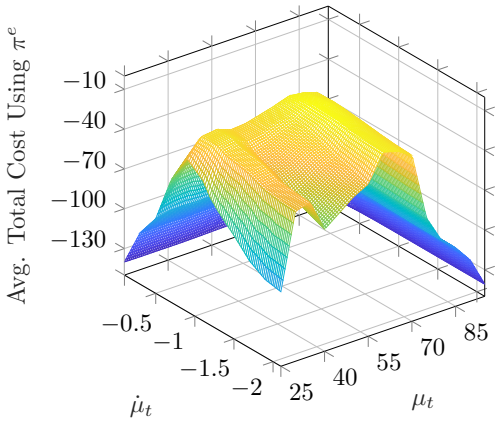
▷ Planning duration $T = 100$ years.
▷ Number of structural elements $E = 1000$.

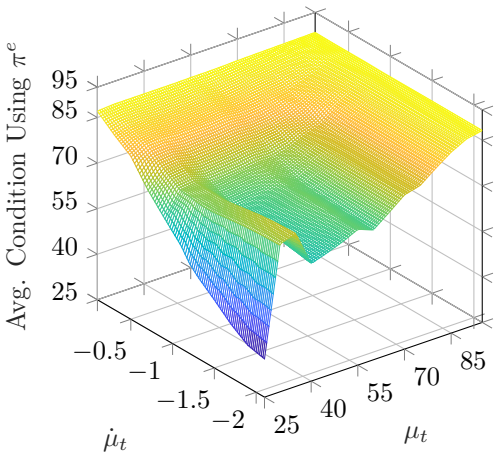
Grid Search Results

Grid Search Results

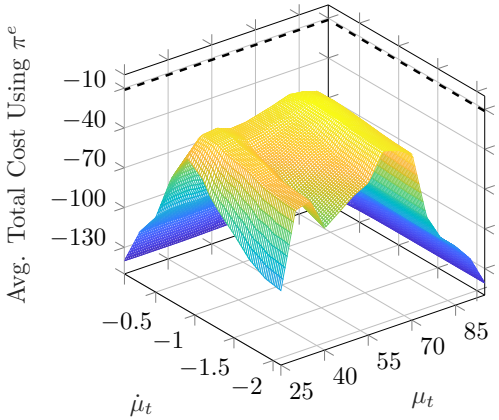


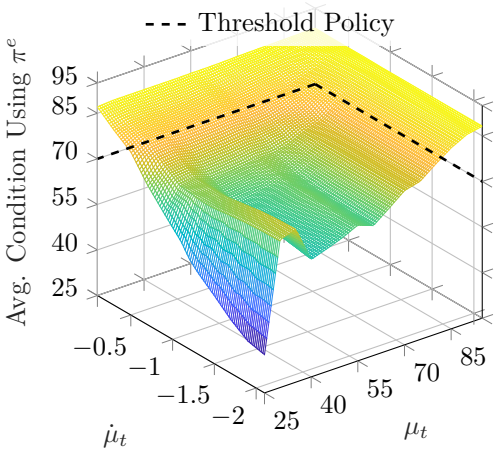
Grid Search Results





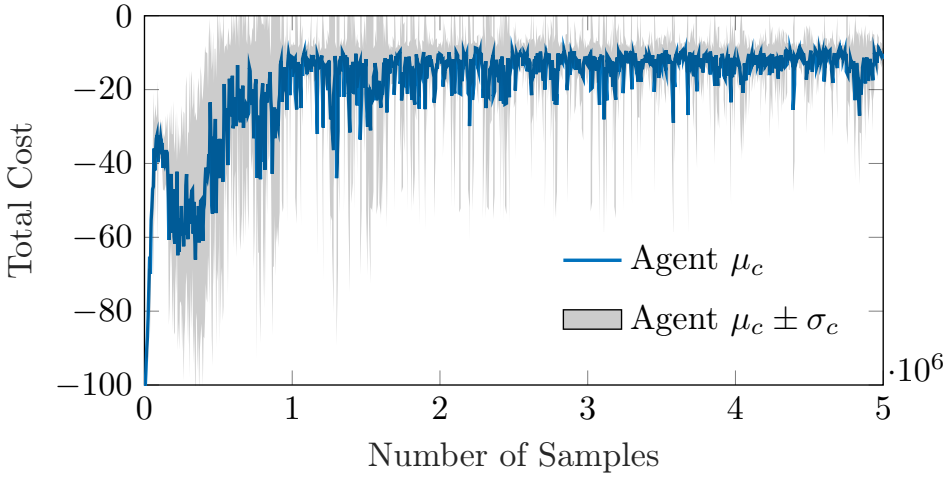
Grid Search Results





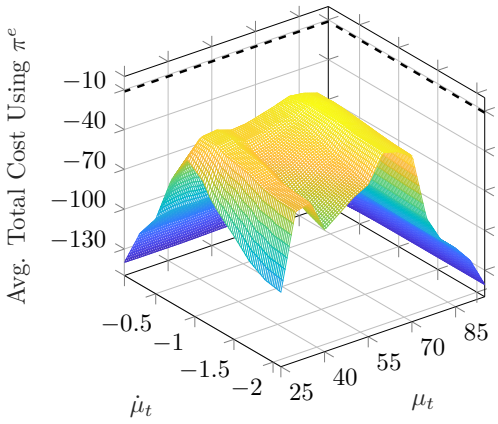
RL Results

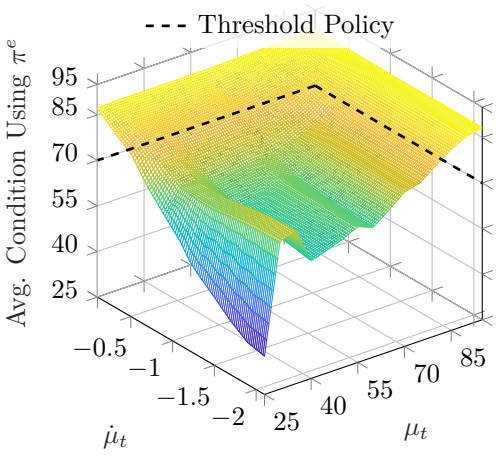
RL Results


0
−20
−40
−60
−80
−100
Total Cost
0
1
2
3
4
5
$\cdot 10^6$
Number of Samples
Agent μ_c
Agent $\mu_c \pm \sigma_c$

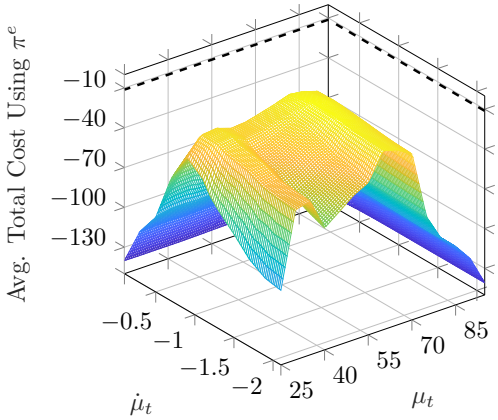
RL vs. Grid Search (Test Set)

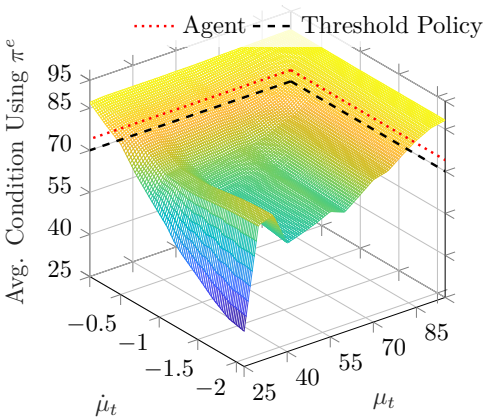
RL vs. Grid Search (Test Set)



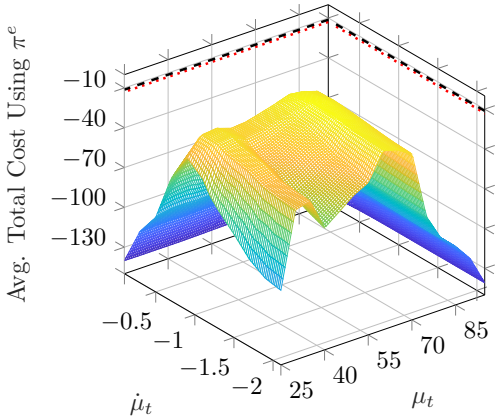


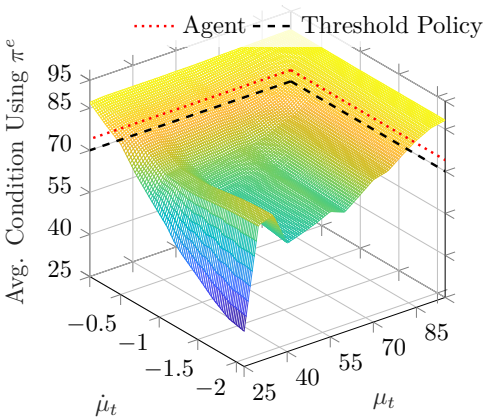
RL vs. Grid Search (Test Set)





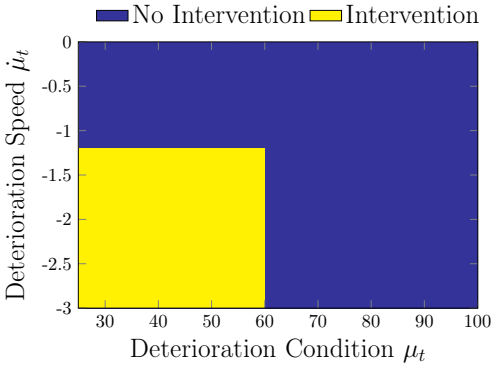
RL vs. Grid Search (Test Set)



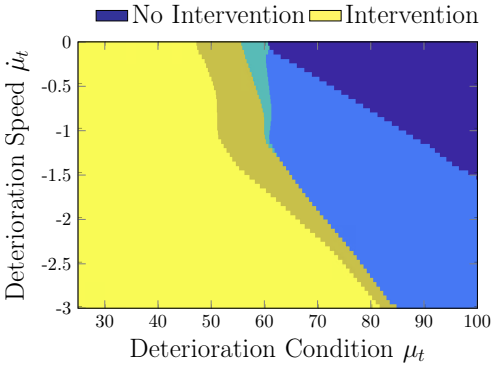


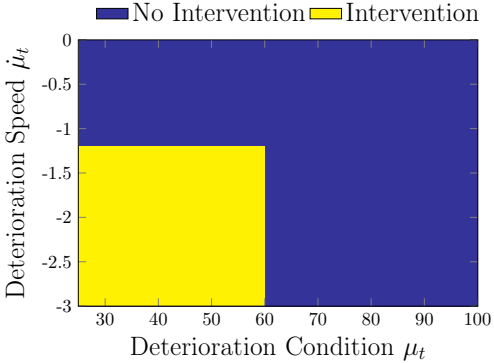
RL vs. GS: Policy

RL vs. GS: Policy

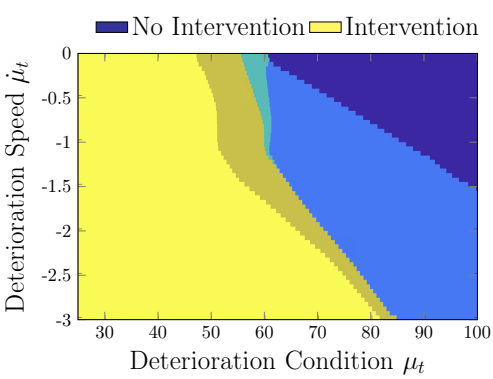


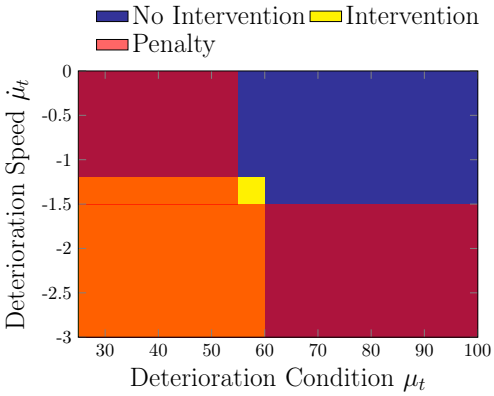
RL vs. GS: Policy



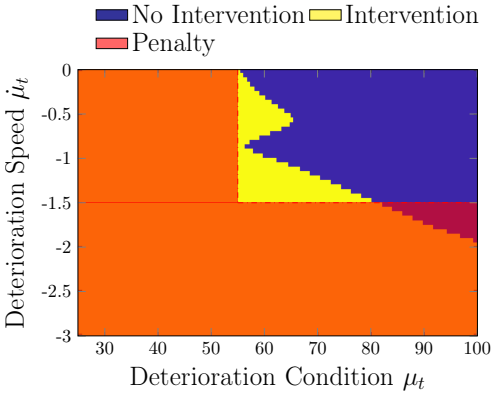


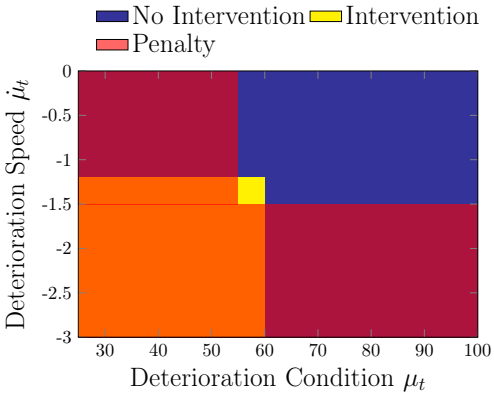
RL vs. GS: Policy





RL vs. GS: Policy





Bridge Maintenance

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b$

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1},$

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1},$

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau,$

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
- ▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

▷ Beams $\times 15$.

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

▷ Beams $\times 15$.
▷ Front walls $\times 2$.

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

▷ Beams $\times 15$.
▷ Front walls $\times 2$.
▷ Slabs $\times 3$.

Bridge Maintenance

- ▷ Problem: Find a maintenance policy for a bridge.
- ▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
- ▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

- ▷ Beams $\times 15$.
- ▷ Front walls $\times 2$.
- ▷ Slabs $\times 3$.
- ▷ Guardrail $\times 2$.

Bridge Maintenance

▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

▷ Beams $\times 15$.
▷ Front walls $\times 2$.
▷ Slabs $\times 3$.
▷ Guardrail $\times 2$.
▷ Wing walls $\times 4$.

Bridge Maintenance

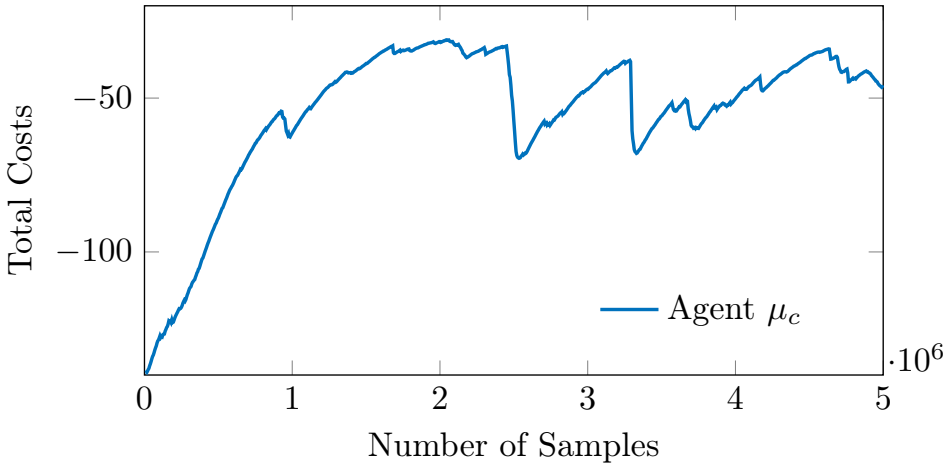
▷ Problem: Find a maintenance policy for a bridge.
▷ Environment: Bridge-level $\boldsymbol{s}_t^b = [\mu_{t|t-1}, \dot{\mu}_{t|t-1}, \delta\tau, var[\boldsymbol{s}_t^c]]$.
▷ Actions: $\mathcal{A}^b = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Bridge Components:

- ▷ Beams $\times 15$.
- ▷ Front walls $\times 2$.
- ▷ Slabs $\times 3$.
- ▷ Guardrail $\times 2$.
- ▷ Wing walls $\times 4$.
- ▷ Pavements $\times 3$.

RL Results (Bridge)

RL Results (Bridge)



Network Maintenance

Network Maintenance

▷ Problem: Find a maintenance policy for a small network.

Network Maintenance

- ▷ Problem: Find a maintenance policy for a small network.
- ▷ Environment: Network-level $\boldsymbol{s}_t^n$

Network Maintenance

- ▷ Problem: Find a maintenance policy for a small network.
- ▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1},$

Network Maintenance

▷ Problem: Find a maintenance policy for a small network.
▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1},$

Network Maintenance

- ▷ Problem: Find a maintenance policy for a small network.
- ▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1}, var[\boldsymbol{s}_t^b]]$.

Network Maintenance

- ▷ Problem: Find a maintenance policy for a small network.
- ▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1}, var[\boldsymbol{s}_t^b]]$.
- ▷ Actions: $\mathcal{A}^n = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Network Maintenance

▷ Problem: Find a maintenance policy for a small network.
▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1}, var[\boldsymbol{s}_t^b]]$.
▷ Actions: $\mathcal{A}^n = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Network Components:

Network Maintenance

▷ Problem: Find a maintenance policy for a small network.
▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1}, var[\boldsymbol{s}_t^b]]$.
▷ Actions: $\mathcal{A}^n = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Network Components:

▷ Bridges $\times 5$.

Network Maintenance

▷ Problem: Find a maintenance policy for a small network.
▷ Environment: Network-level $\boldsymbol{s}_t^n = [\boldsymbol{\mu}_{t|t-1}, \dot{\boldsymbol{\mu}}_{t|t-1}, var[\boldsymbol{s}_t^b]]$.
▷ Actions: $\mathcal{A}^n = \{a_0 : \text{do nothing}, a_1 : \text{maintain}\}$.

Network Components:

- ▷ Bridges $\times 5$.
- ▷ Network connectivity (Bridges $\times 2$).

Limitations

Limitations

- Sub-optimal policies due to:

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.
- Unaddressed challenges.

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.
- Unaddressed challenges.
  - ▷ Time varying costs: g(deterioration level, element size).

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.
- Unaddressed challenges.
  - ▷ Time varying costs: g(deterioration level, element size).
  - ▷ Factoring duration of service stoppage.

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.
- Unaddressed challenges.
  - ▷ Time varying costs: g(deterioration level, element size).
  - ▷ Factoring duration of service stoppage.
  - ▷ Network connectivity & re-routing costs.

Limitations

- Sub-optimal policies due to:
  - ▷ Sequential training in the hierarchy.
  - ▷ Solving a partially observable environment.
- Unaddressed challenges.
  - ▷ Time varying costs: g(deterioration level, element size).
  - ▷ Factoring duration of service stoppage.
  - ▷ Network connectivity & re-routing costs.
- Extend the application to a high number of structures.

Conclusions & Future Work

Conclusions & Future Work

- Hierarchical formulation is suitable for network-scale planning.

Conclusions & Future Work

- Hierarchical formulation is suitable for network-scale planning.
- Multi-agent RL solvers can potentially extend the scalability of HRL.

Conclusions & Future Work

- Hierarchical formulation is suitable for network-scale planning.
- Multi-agent RL solvers can potentially extend the scalability of HRL.
- In real world applications, RL is NOT perfect → partial control,

Conclusions & Future Work

- Hierarchical formulation is suitable for network-scale planning.
- Multi-agent RL solvers can potentially extend the scalability of HRL.
- In real world applications, RL is NOT perfect → partial control,
  - ▷ Supervised learning.

Conclusions & Future Work

- Hierarchical formulation is suitable for network-scale planning.
- Multi-agent RL solvers can potentially extend the scalability of HRL.
- In real world applications, RL is NOT perfect → partial control,
  - ▷ Supervised learning.
  - ▷ Decision verification systems.