Network-Scale Maintenance Planning for Infrastructures Using Reinforcement Learning (Formulation)

Zachary Hamida Postdoc
James-A. Goulet Professeur

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

August 6, 2021

Funding:
Transportation Ministry of Quebec (MTQ)

Outline

Recap & Definitions

Visual Inspections (VI): Network-scale monitoring technique

Source: Google images

Recap & Definitions

Visual Inspections (VI): Network-scale monitoring technique

Network:

Source: Google images

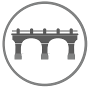
Recap & Definitions

Visual Inspections (VI): Network-scale monitoring technique

Network:

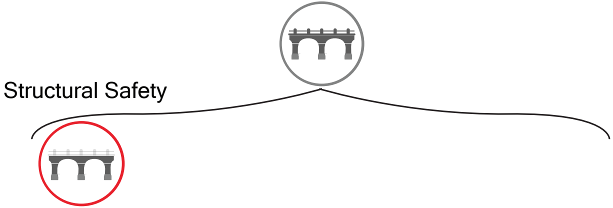
Source: Google images

Data Hierarchy



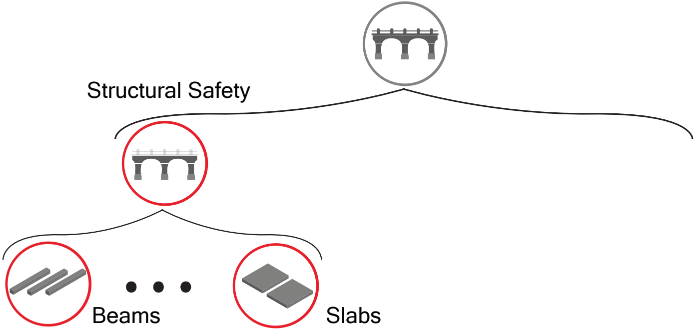
Source: MTQ, Manual of Inspection

Data Hierarchy



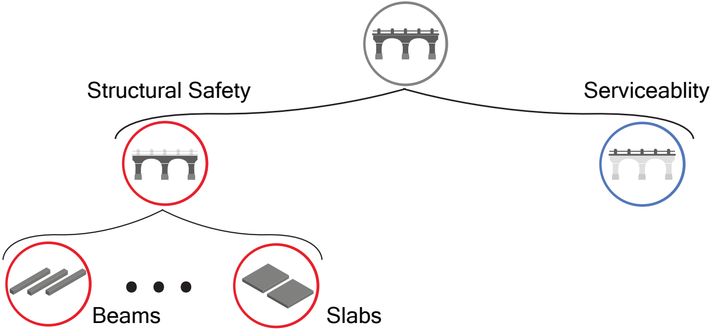
Source: MTQ, Manual of Inspection

Data Hierarchy



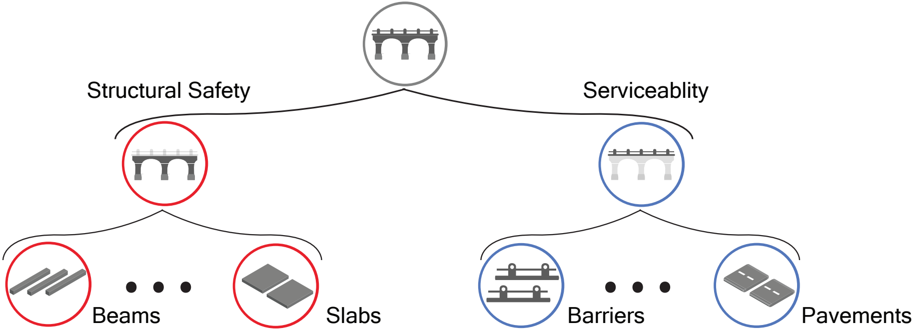
Source: MTQ, Manual of Inspection

Data Hierarchy



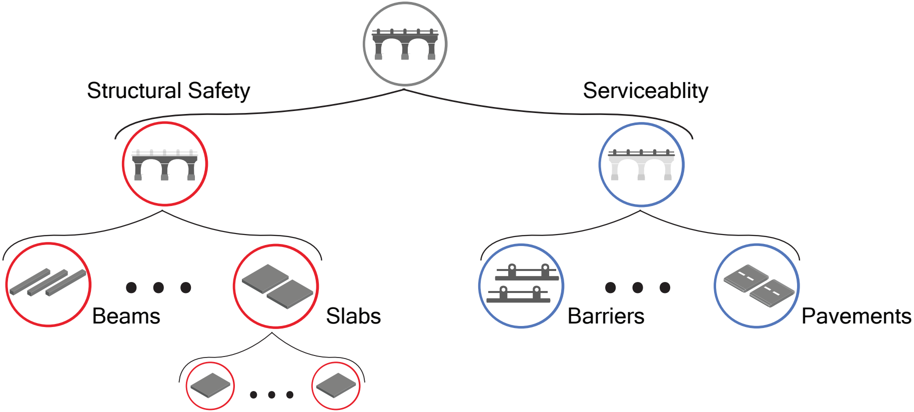
Source: MTQ, Manual of Inspection

Data Hierarchy



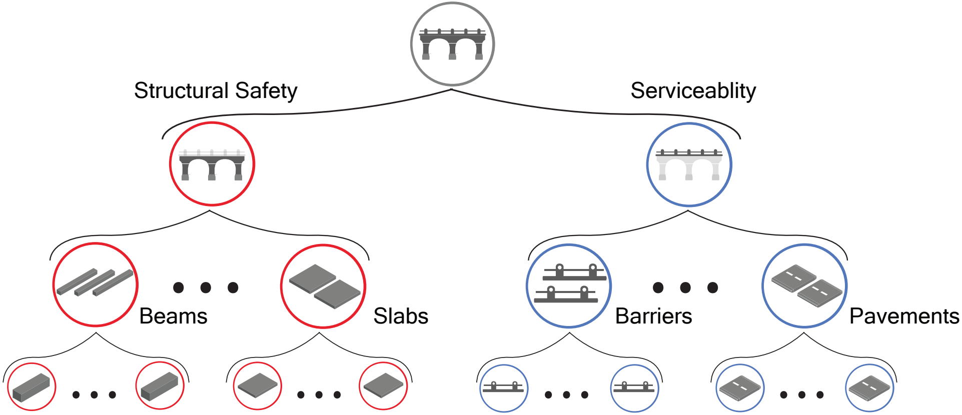
Source: MTQ, Manual of Inspection

Data Hierarchy



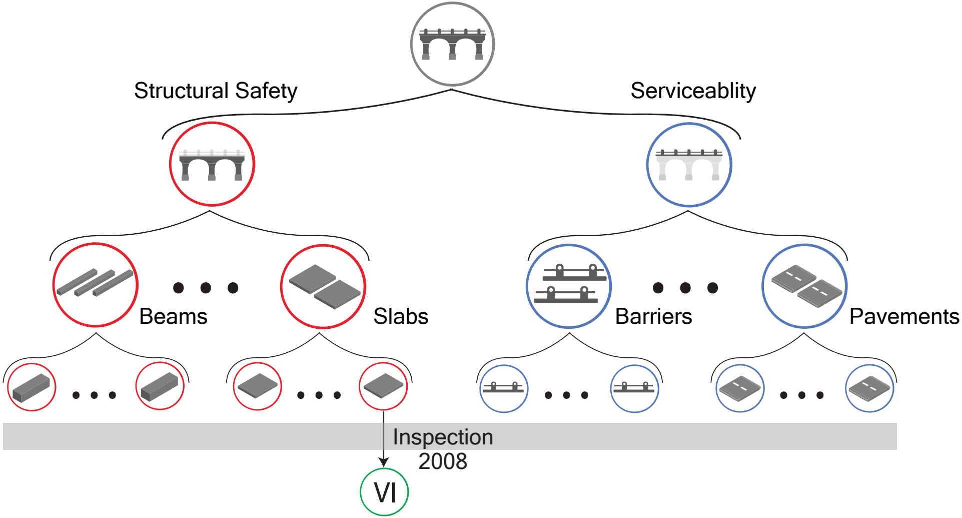
Source: MTQ, Manual of Inspection

Data Hierarchy



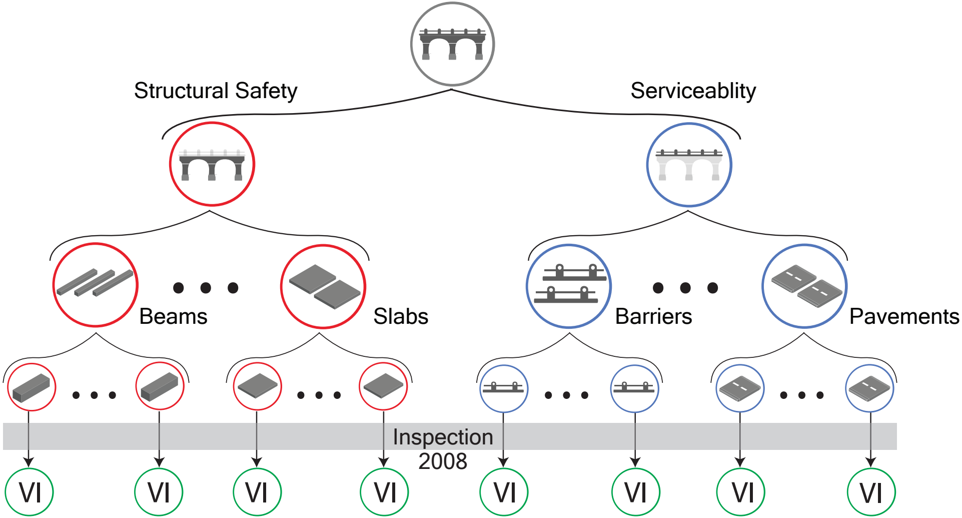
Source: MTQ, Manual of Inspection

Data Hierarchy



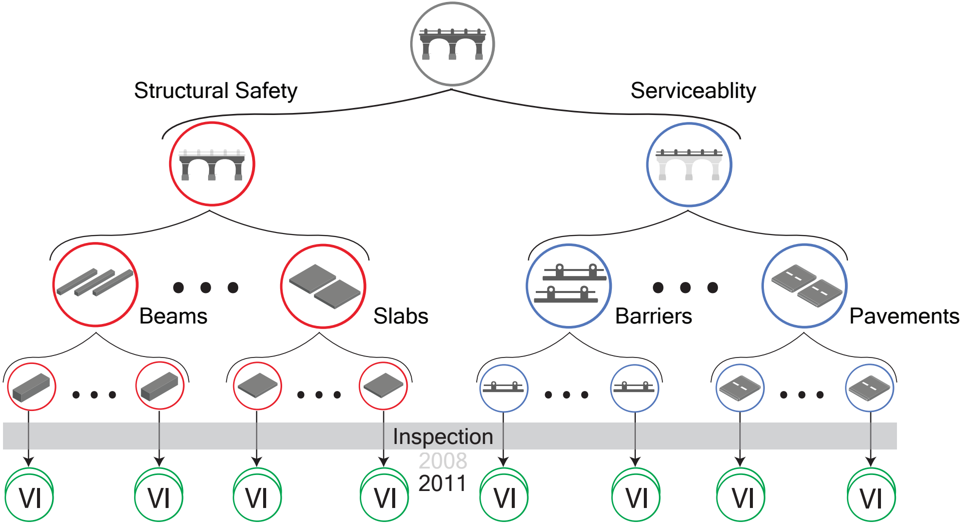
Source: MTQ, Manual of Inspection

Data Hierarchy



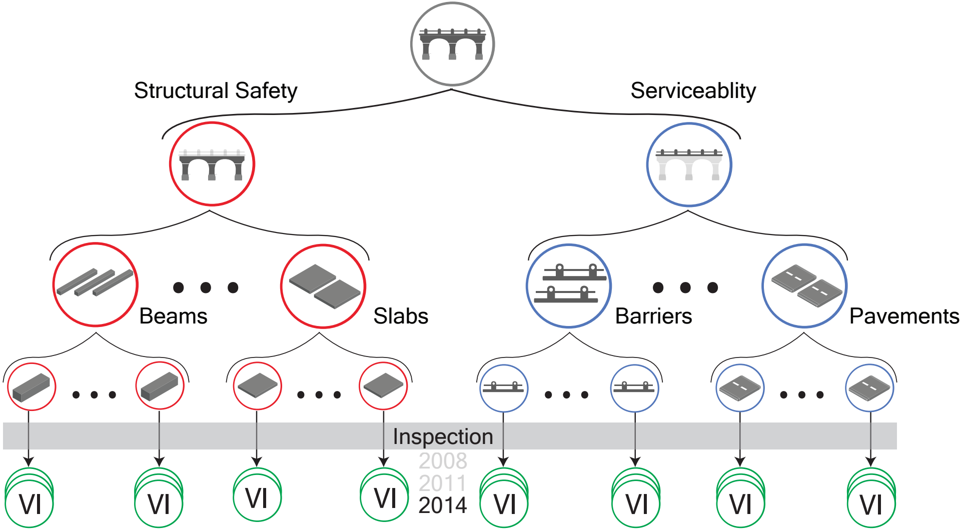
Source: MTQ, Manual of Inspection

Data Hierarchy



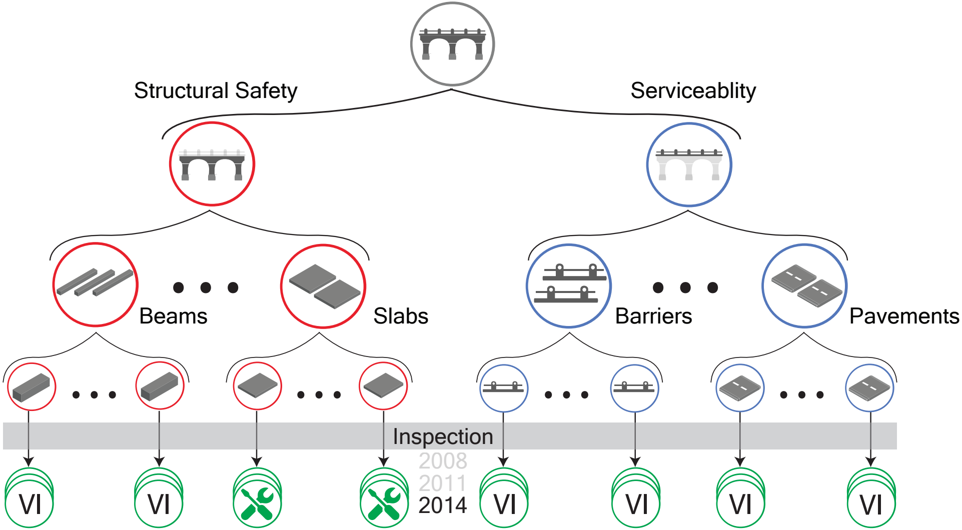
Source: MTQ, Manual of Inspection

Data Hierarchy



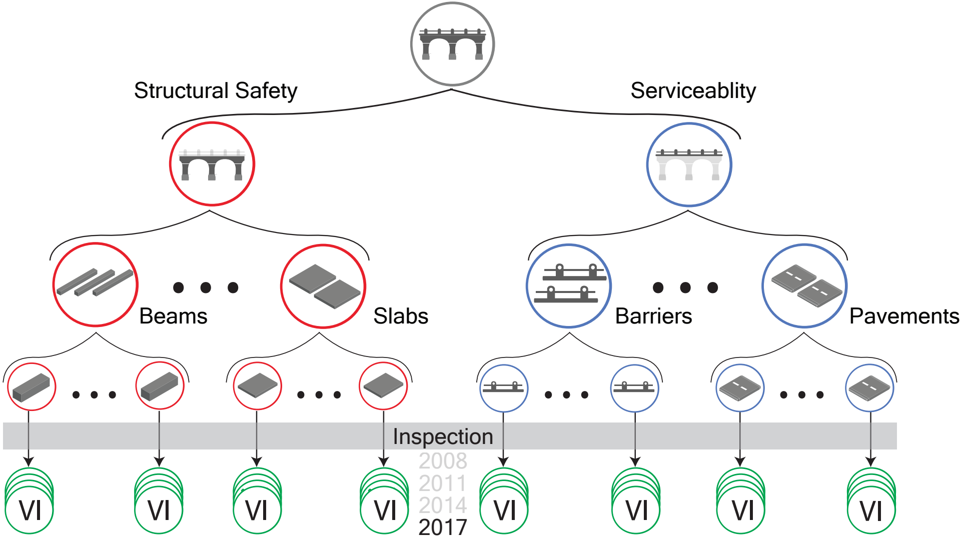
Source: MTQ, Manual of Inspection

Data Hierarchy



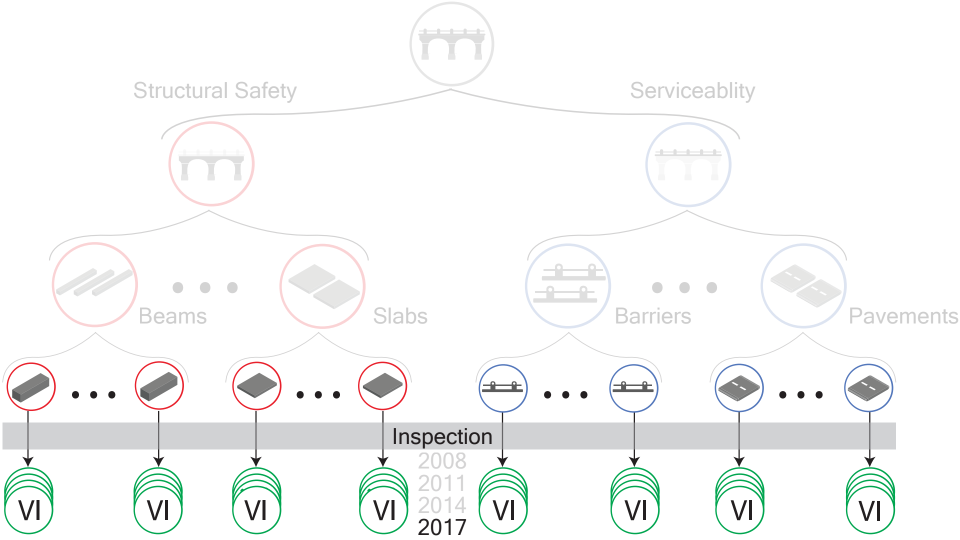
Source: MTQ, Manual of Inspection

Data Hierarchy



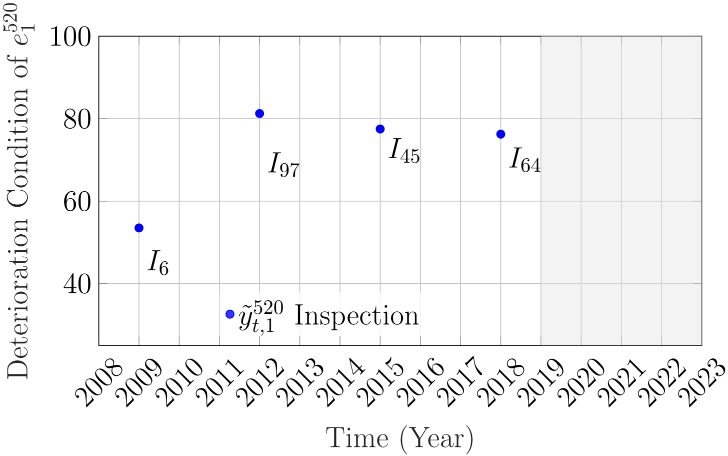
Source: MTQ, Manual of Inspection

Data Hierarchy

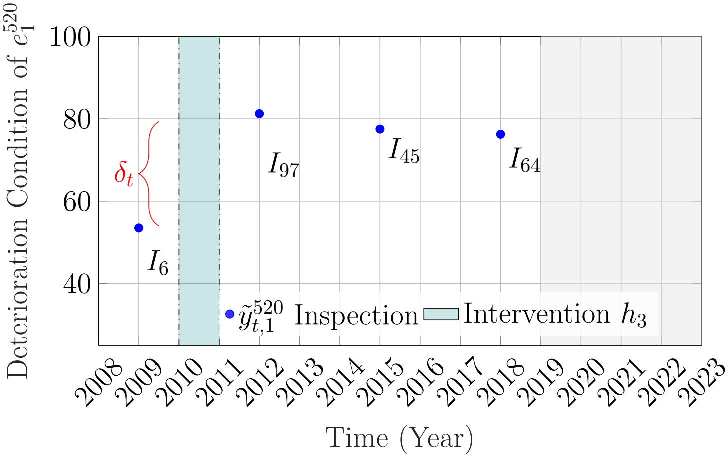


Source: MTQ, Manual of Inspection

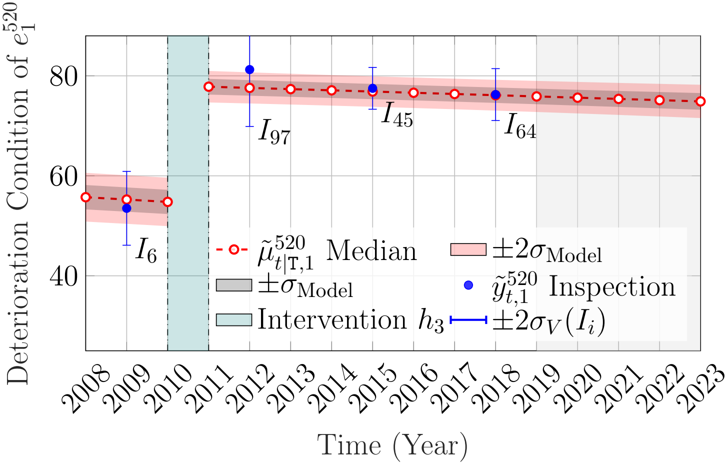
Example of visual inspection data with an intervention - Repairs

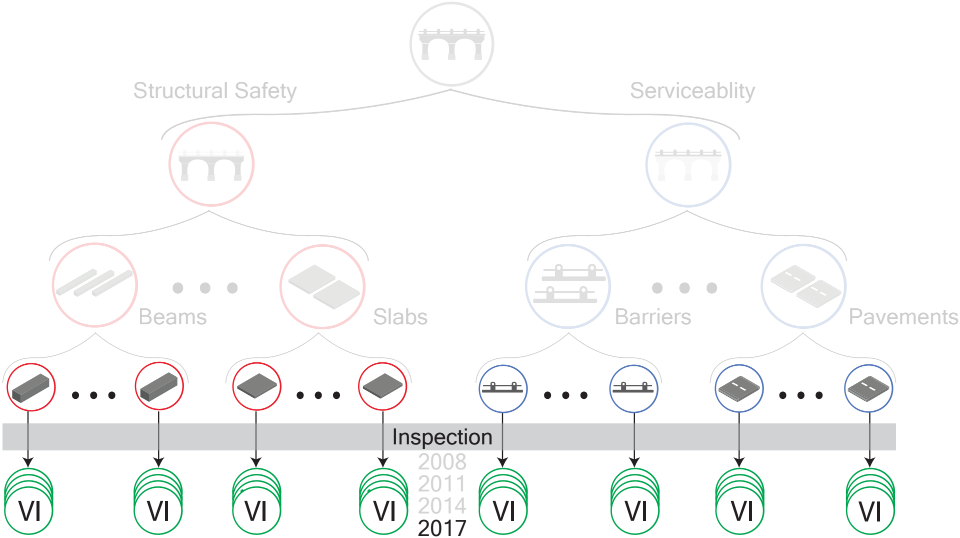


Example of visual inspection data with an intervention - Repairs

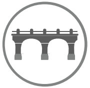


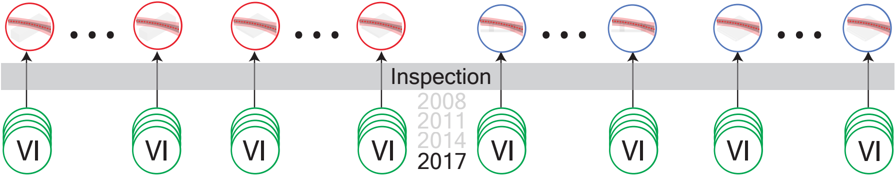
Example of visual inspection data with an intervention - Repairs

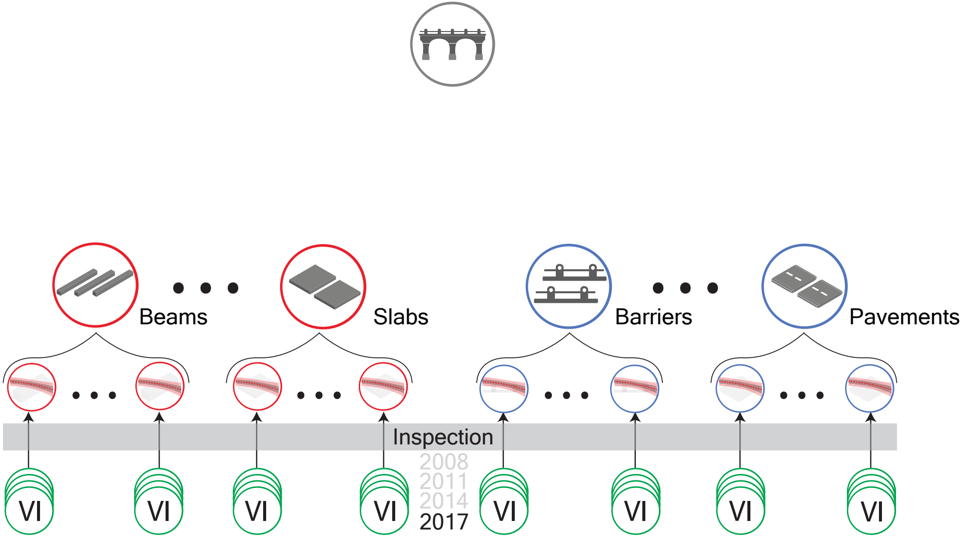


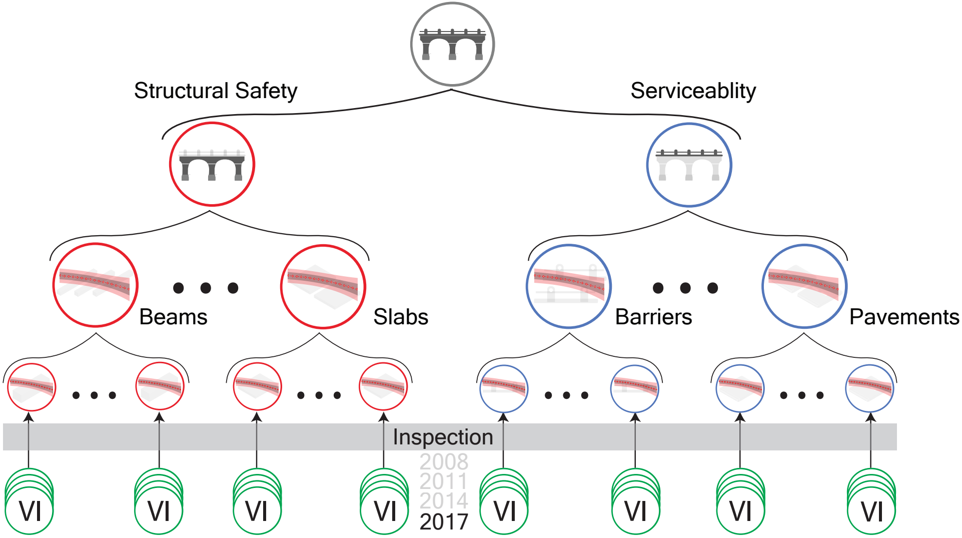

Structural Safety
Serviceablity
Beams
Slabs
Barriers
Pavements
Inspection
2008
2011
2014
2017
VI

Recap & Definitions






Beams
Slabs
Barriers
Pavements
Inspection
2008
2011
2014
2017
VI


Structural Safety
Serviceablity
Beams
Slabs
Barriers
Pavements
Inspection
2008
2011
2014
2017
VI

Project Progress

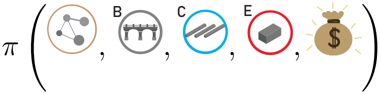
☑ Model the deterioration based on visual inspections.

Project Progress

- [x] Model the deterioration based on visual inspections.
- [x] Quantify the effect of interventions.
- [ ] Use the framework for planning interventions (decision making).

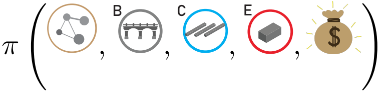
Project Progress

Find a policy for maintenance:



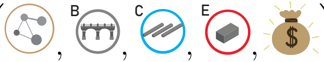
Project Progress

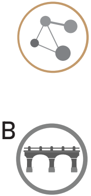
Find a policy for maintenance:



Project Progress

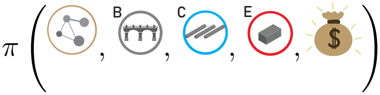
Find a policy for maintenance:

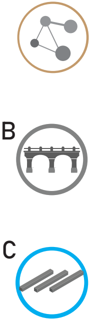
π ()



Project Progress

Find a policy for maintenance:

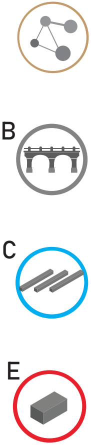




Project Progress

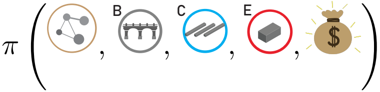
Find a policy for maintenance:

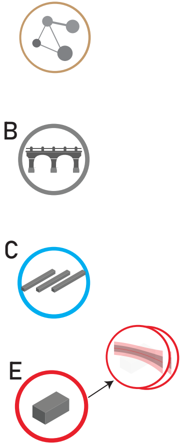
$\pi\Big($  $\Big)$



Project Progress

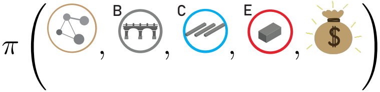
Find a policy for maintenance:

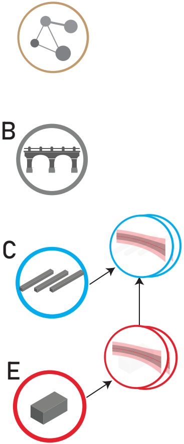




Project Progress

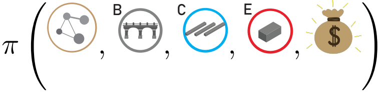
Find a policy for maintenance:

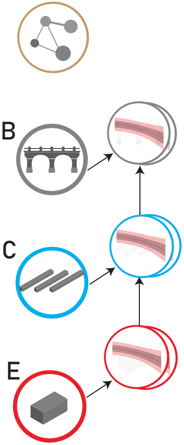




Project Progress

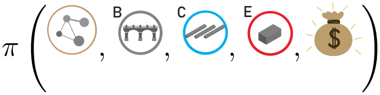
Find a policy for maintenance:

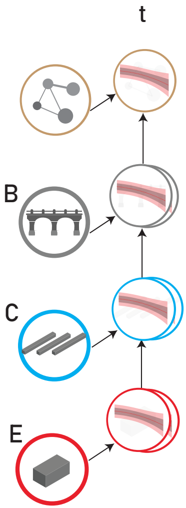




Project Progress

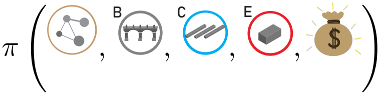
Find a policy for maintenance:

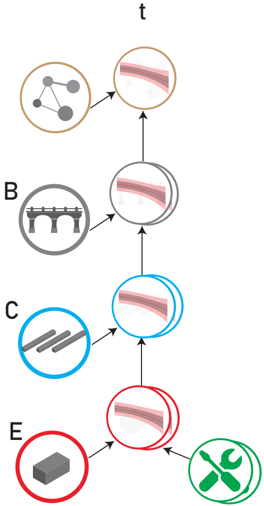




Project Progress

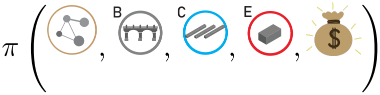
Find a policy for maintenance:

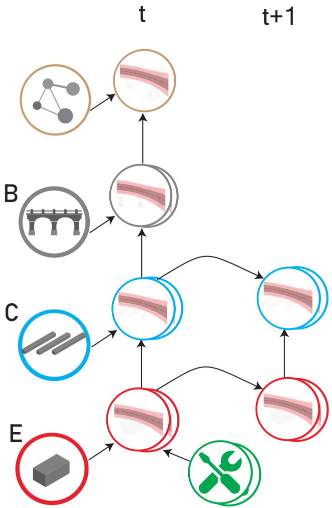




Project Progress

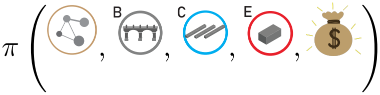
Find a policy for maintenance:

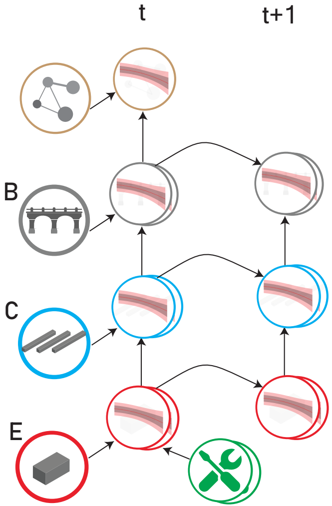




Project Progress

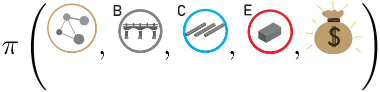
Find a policy for maintenance:

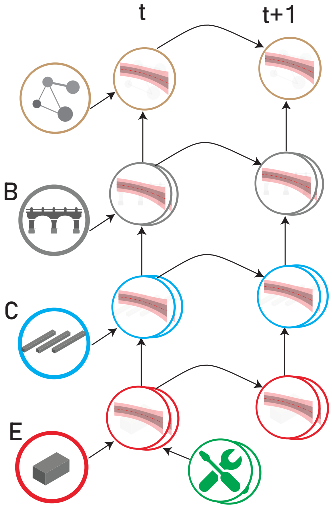




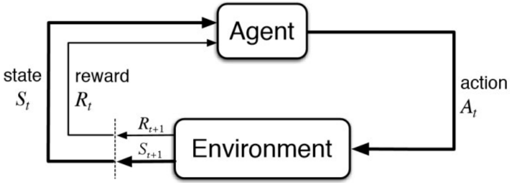
Project Progress

Find a policy for maintenance:



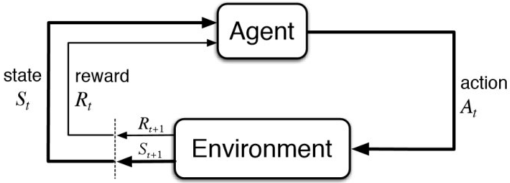


Reinforcement Learning



Source: Google Images

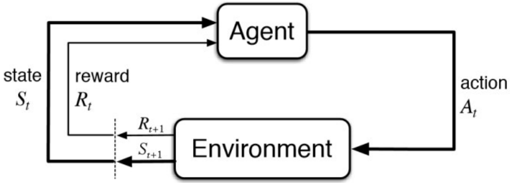
Reinforcement Learning



- Agent: decision maker.

Source: Google Images

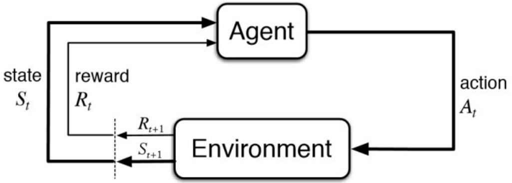
Reinforcement Learning



- Agent: decision maker.
- Environment: e.g., video game.

Source: Google Images

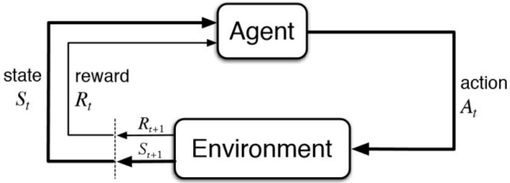
Reinforcement Learning



- ▷ Agent: decision maker.
- ▷ Environment: e.g., video game.
- ▷ Actions: e.g., key press or movement.

Source: Google Images

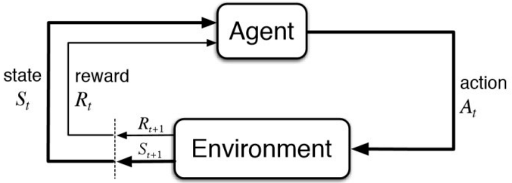
Reinforcement Learning



- ▷ Agent: decision maker.
- ▷ Environment: e.g., video game.
- ▷ Actions: e.g., key press or movement.
- ▷ Rewards: e.g., achieving the goal.

Source: Google Images

Reinforcement Learning

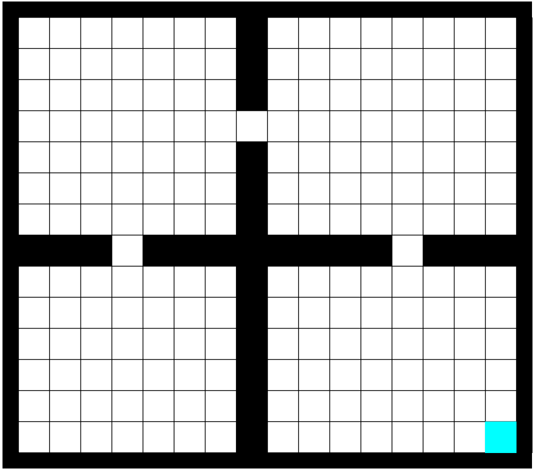


- Agent: decision maker.
- Environment: e.g., video game.
- Actions: e.g., key press or movement.
- Rewards: e.g., achieving the goal.

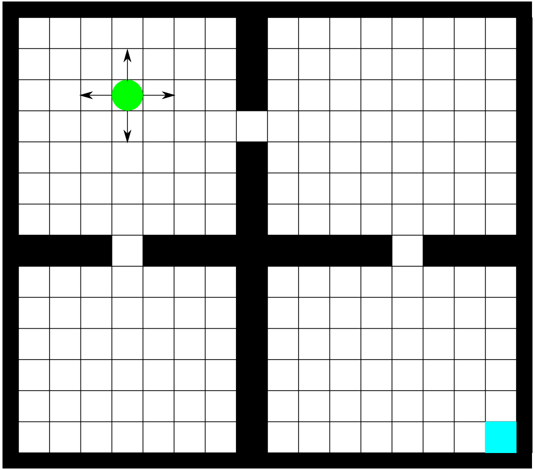
Learn a policy that maximizes the total expected discounted rewards.

Source: Google Images

Example:

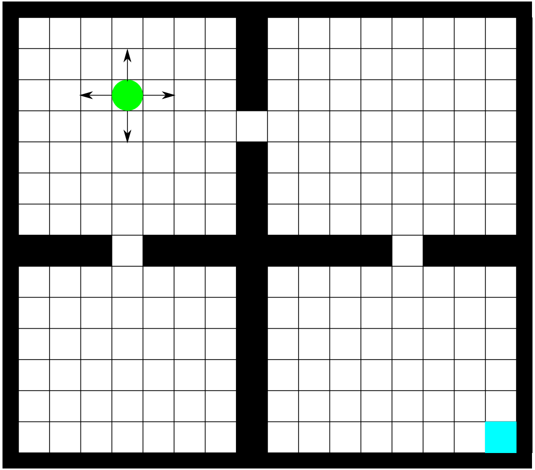


Example:



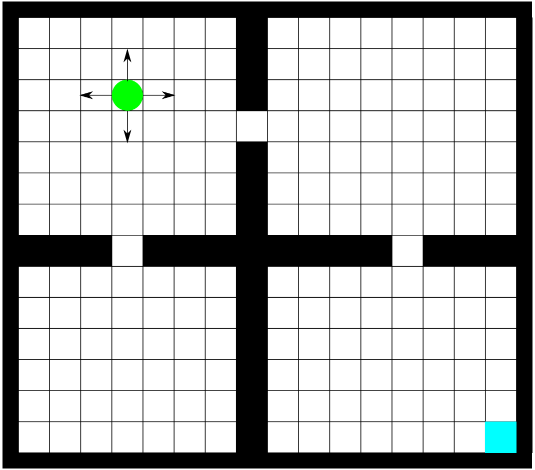
Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.

Example:



Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Learn a policy π based on maximizing a function $Q(s, a)$.

Example:

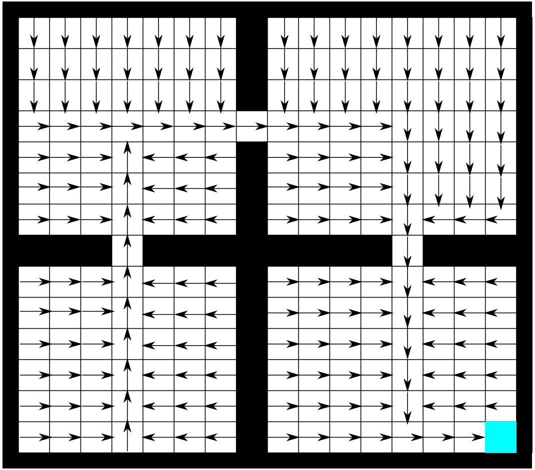


Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Learn a policy π based on maximizing a function $Q(s, a)$.

Relying on the agent experience: (s_t, a_t, r_t, s_{t+1})

Example: $([2, 2], \rightarrow, 0, [3, 2])$.

Example:

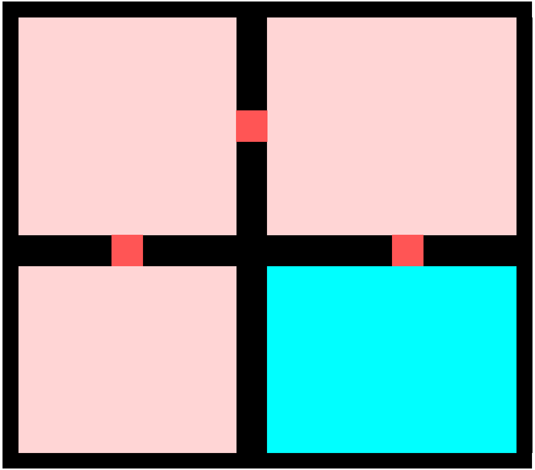


Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Learn a policy π based on maximizing a function $Q(s, a)$.

Relying on the agent experience:
(s_t, a_t, r_t, s_{t+1})

Example: $([2, 2], \rightarrow, 0, [3, 2])$.

Example:

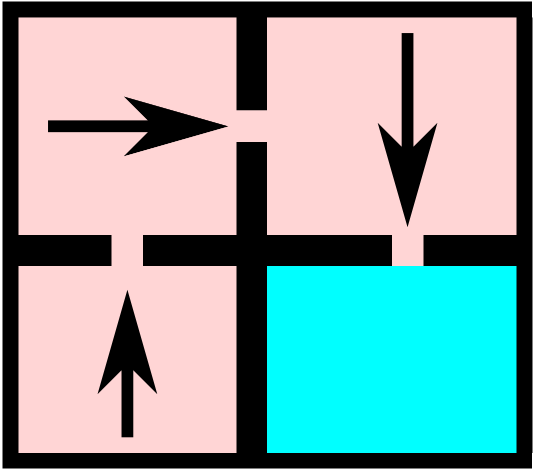


Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Additional components: $g_\omega, I_\omega, \beta_\omega, r_\omega$

Learn a global policy π and local policies π_ω using the same concepts of Q-Learning.

Hierarchical Reinforcement Learning.

Example:

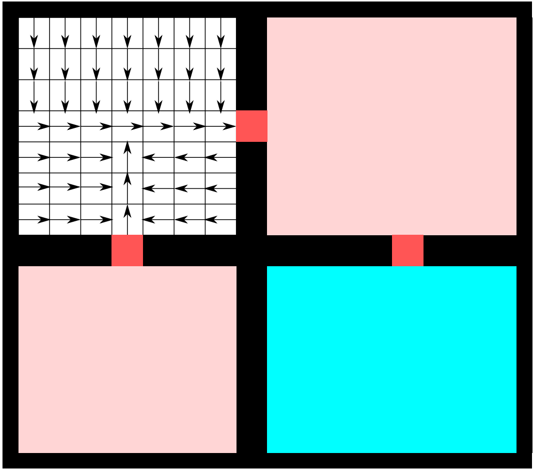


Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Additional components: $g_\omega, l_\omega, \beta_\omega, r_\omega$

Learn a global policy π and local policies π_ω using the same concepts of Q-Learning.

Hierarchical Reinforcement Learning.

Example:

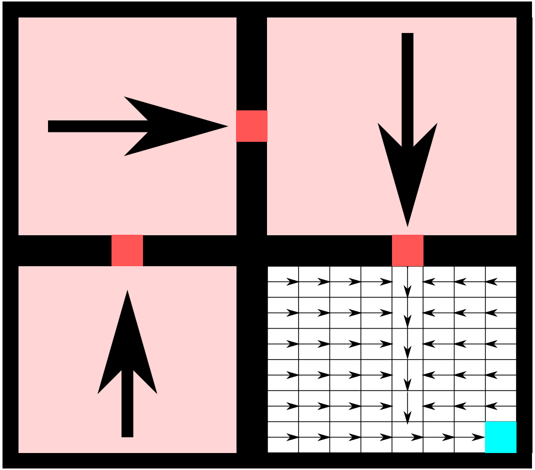


Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Additional components: $g_\omega, l_\omega, \beta_\omega, r_\omega$

Learn a global policy π and local policies π_ω using the same concepts of Q-Learning.

Hierarchical Reinforcement Learning.

Example:

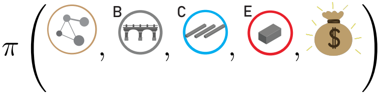


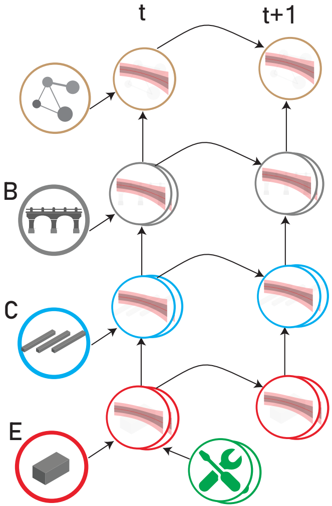
Goal: get to the charging station.
Environment: 4 rooms.
Actions: $\leftarrow, \rightarrow, \uparrow, \downarrow$
Rewards: 1 achieving the goal, 0 otherwise.
Additional components: $g_\omega, l_\omega, \beta_\omega, r_\omega$

Learn a global policy π and local policies π_ω using the same concepts of Q-Learning.

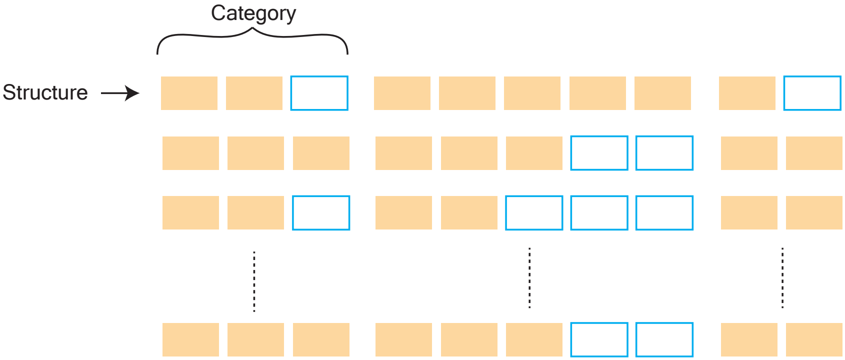
Hierarchical Reinforcement Learning.

Find a policy for maintenance:

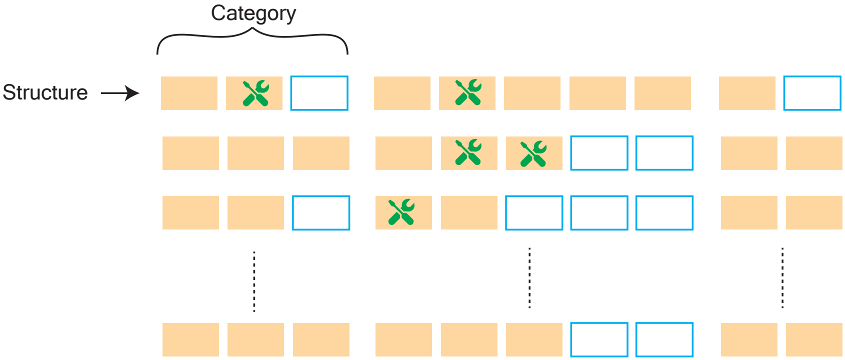




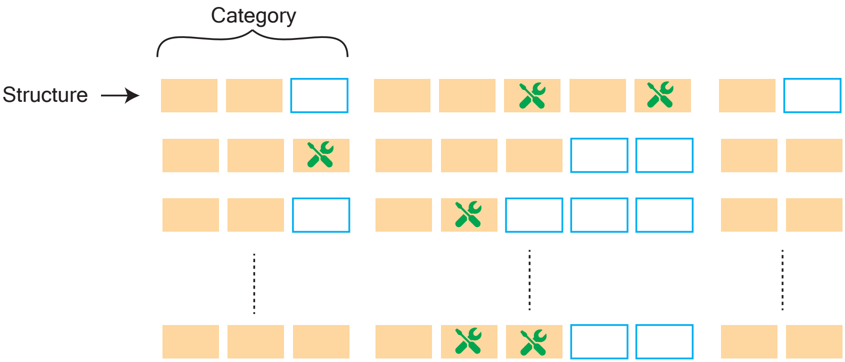
The state of the network at time t:

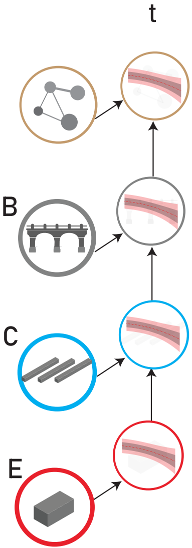


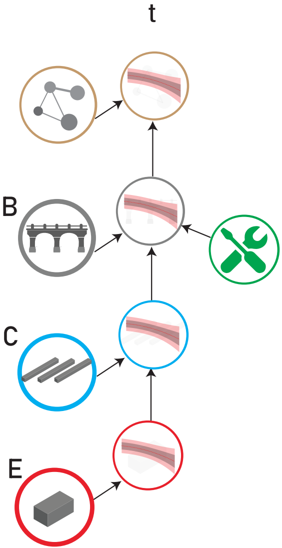
The state of the network at time t:

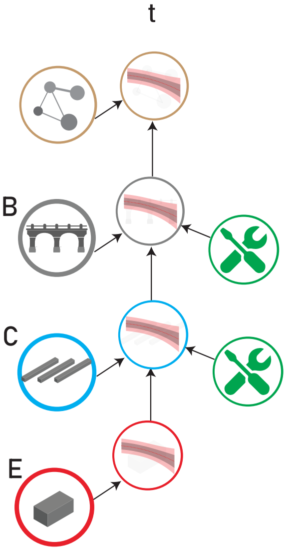


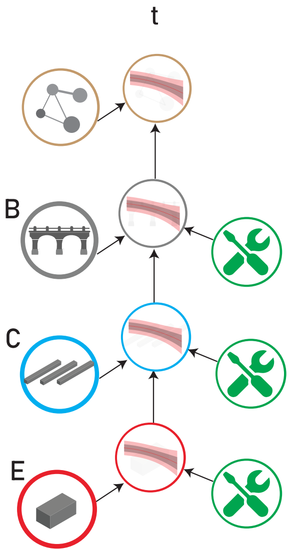
The state of the network at time t:

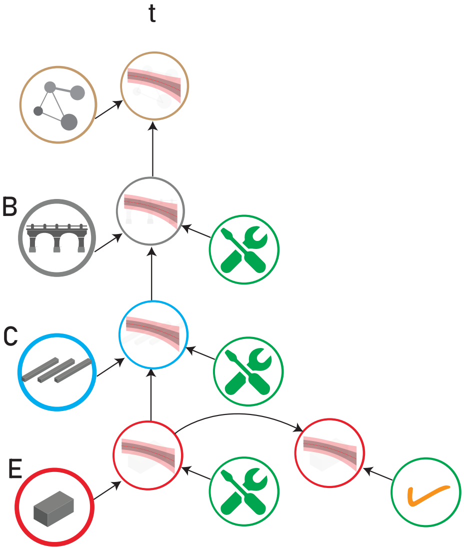


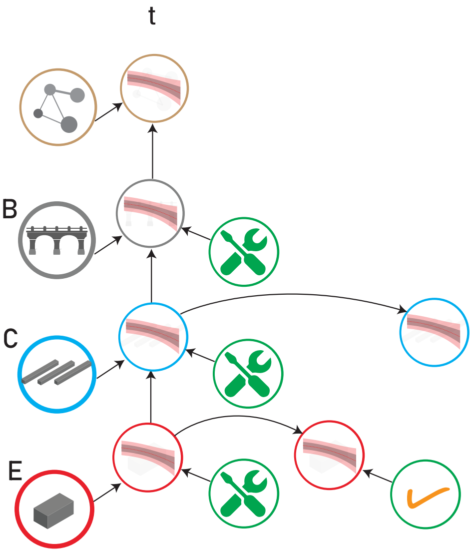

t
B
C
E

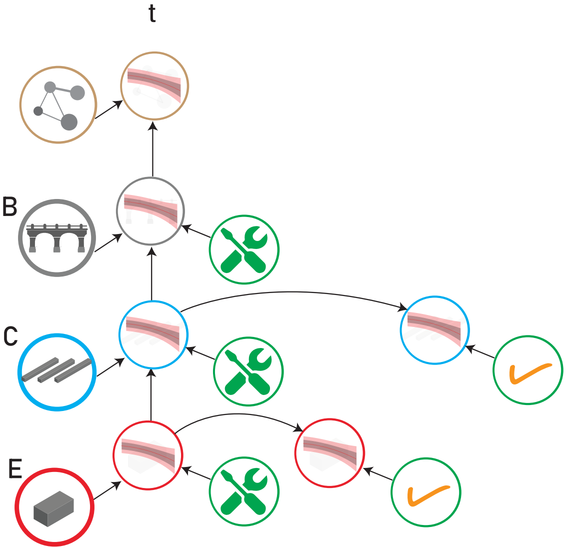

t
B
C
E

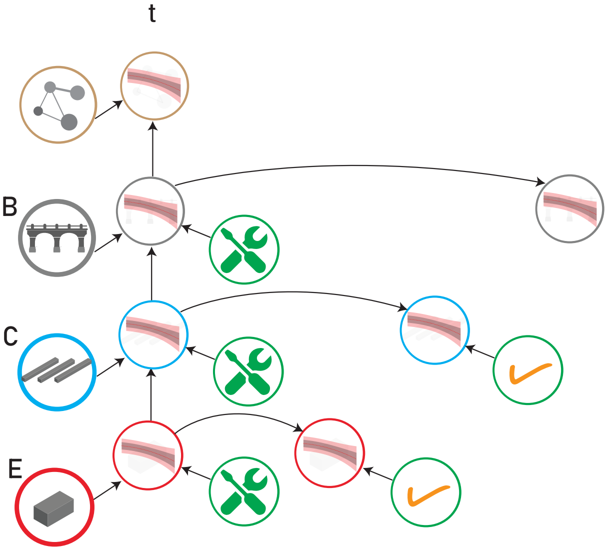

t
B
C
E

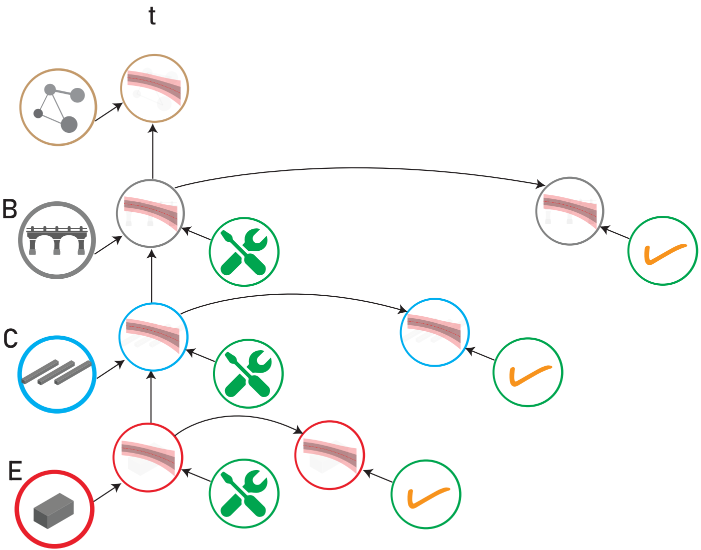

t
B
C
E

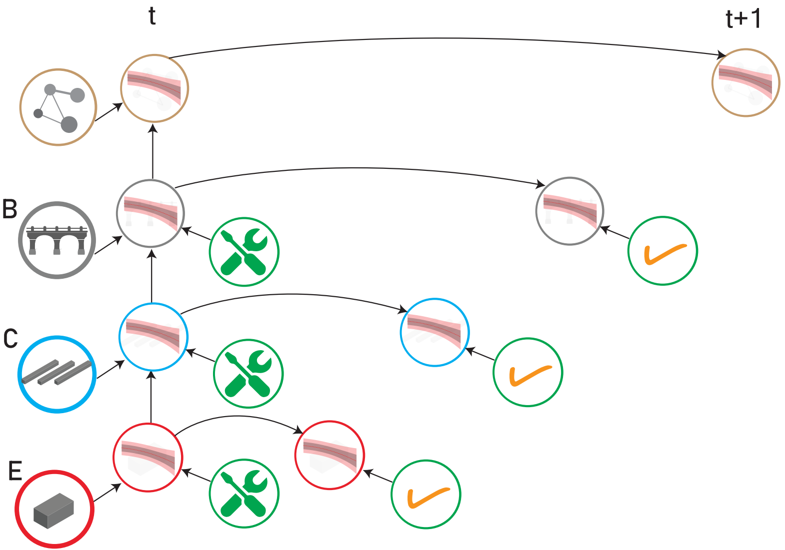

t
B
C
E


t
B
C
E

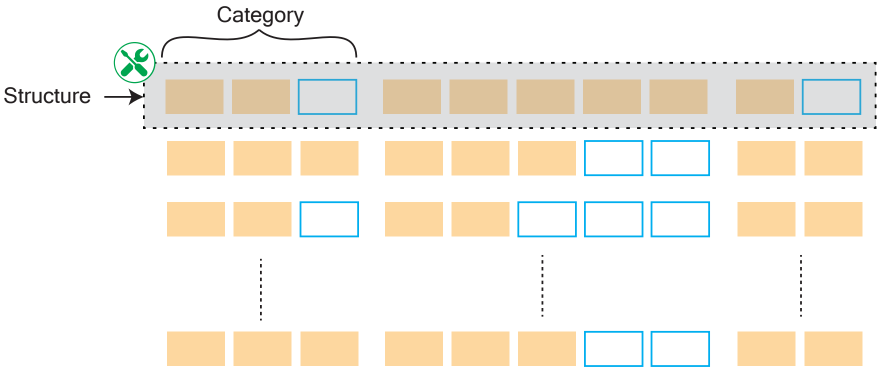

t
B
C
E


t
B
C
E

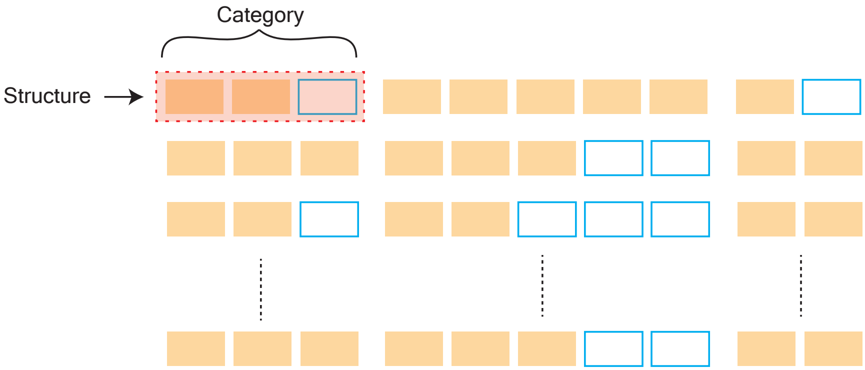

t
B
C
E


t
t+1
B
C
E

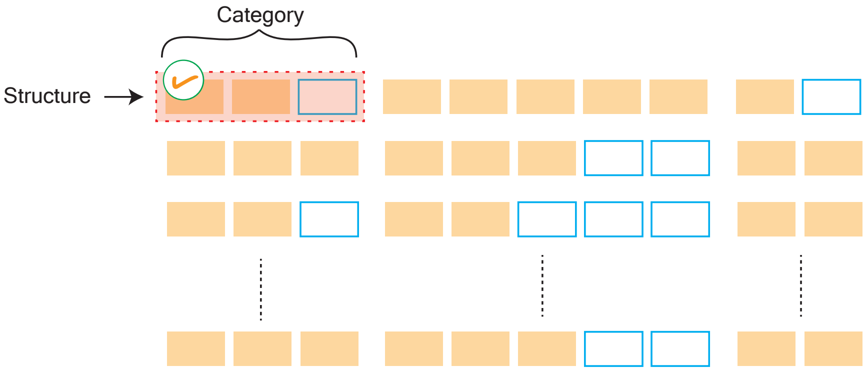
The state of the network at time t:



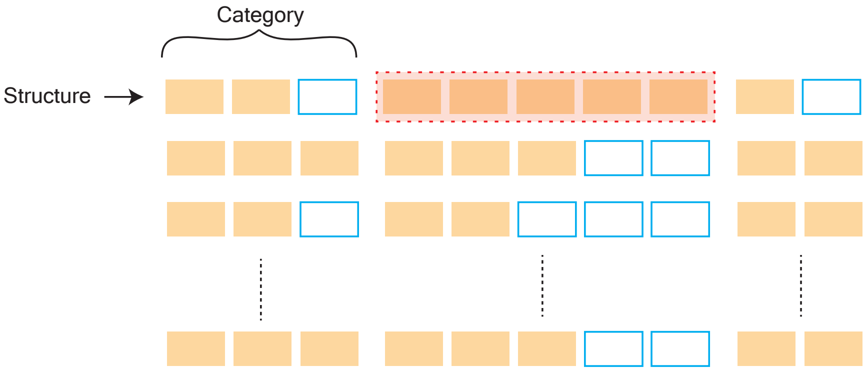
The state of the network at time t:



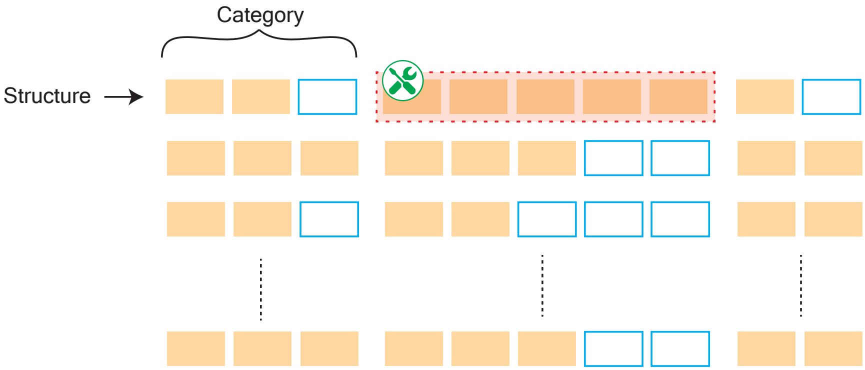
The state of the network at time t:



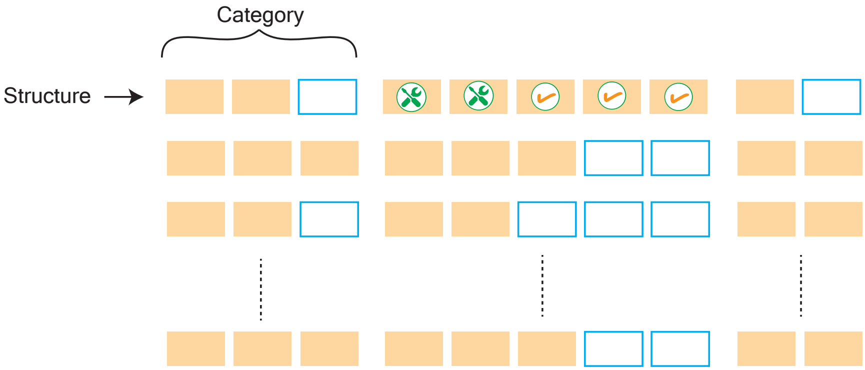
The state of the network at time t:



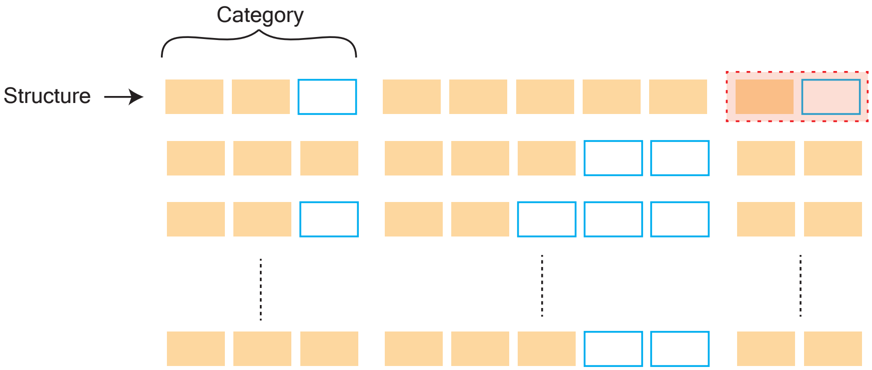
The state of the network at time t:



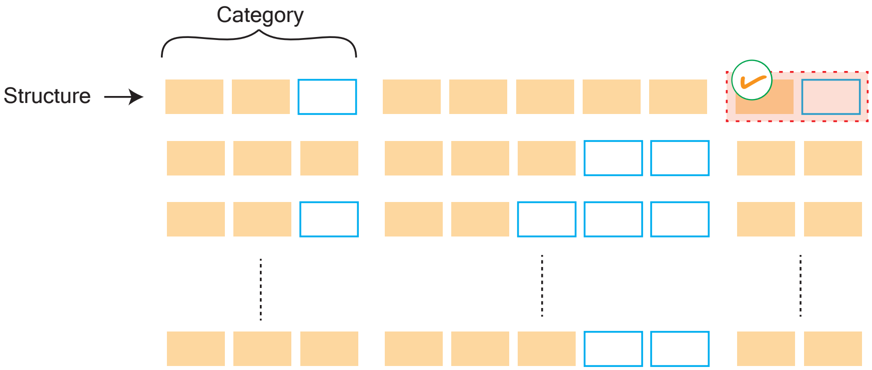
The state of the network at time t:



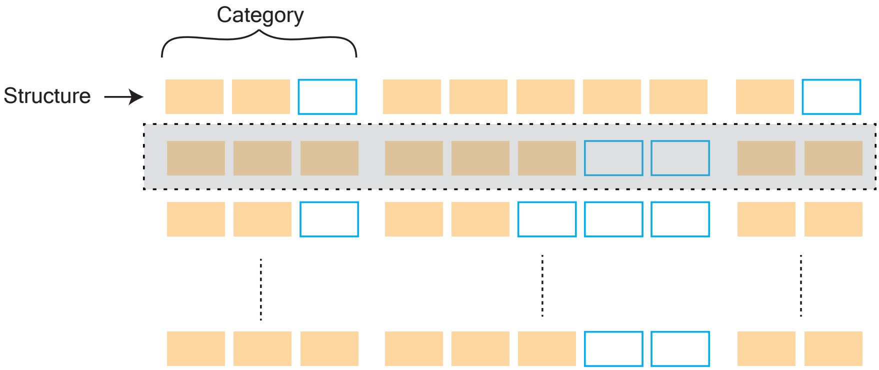
The state of the network at time t:



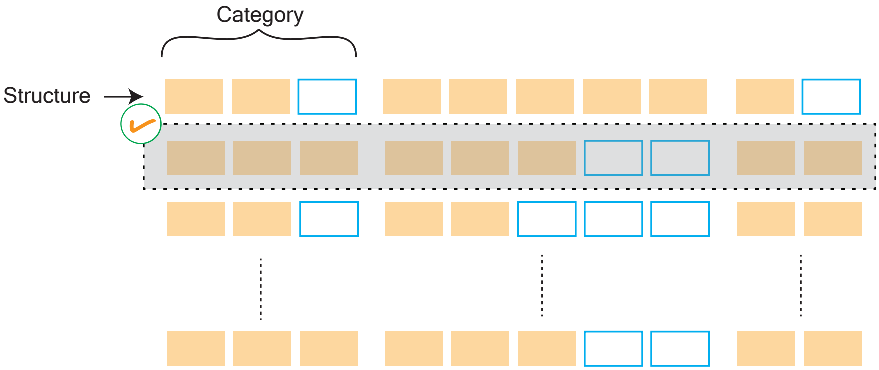
The state of the network at time t:



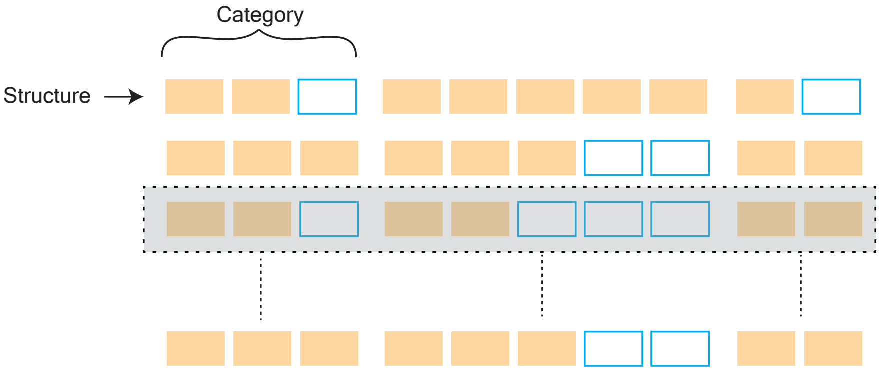
The state of the network at time t:



The state of the network at time t:



The state of the network at time t:



Next Steps

- Improve environment design.

Next Steps

- ▷ Improve environment design.
- ▷ Design the reward functions.

Next Steps

- ▷ Improve environment design.
- ▷ Design the reward functions.
- ▷ Explore policy training options.

Next Steps

- ▷ Improve environment design.
- ▷ Design the reward functions.
- ▷ Explore policy training options.