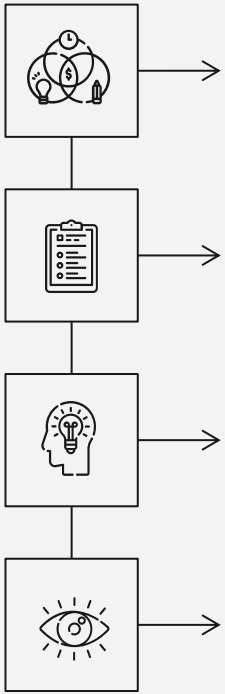
Empowering Epidemic Response:

The Role of Reinforcement Learning in Infectious Disease Control

Mutong LIU (Presenter), Prof. Yang LIU, Prof. Jiming LIU
Department of Computer Science, Hong Kong Baptist University

Contents



01 → Background of Infectious Disease Control

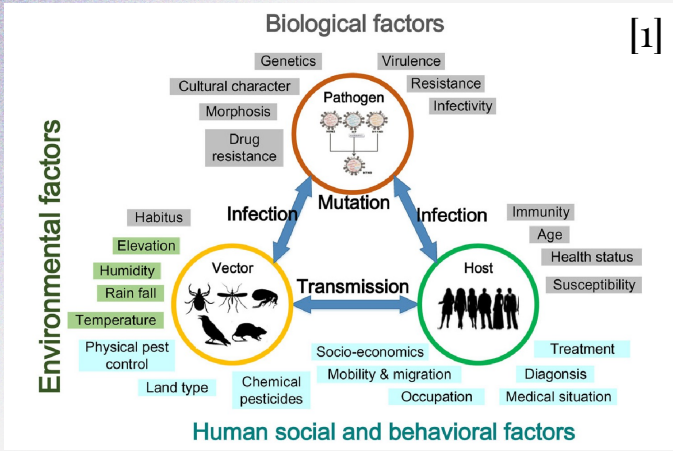
02 → Literature Search Methodology

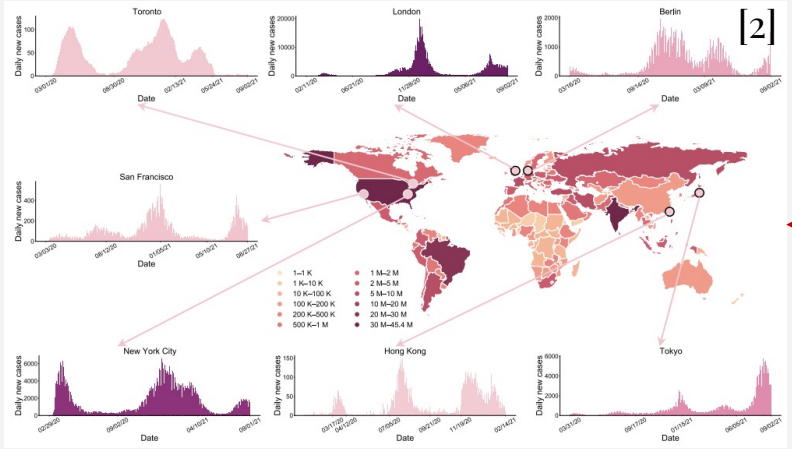
03 → Survey Taxonomy

04 → Conclusion and Future Directions

01
Background of Infectious Disease Control

Disease Spread and Control





Epidemic Control Interventions

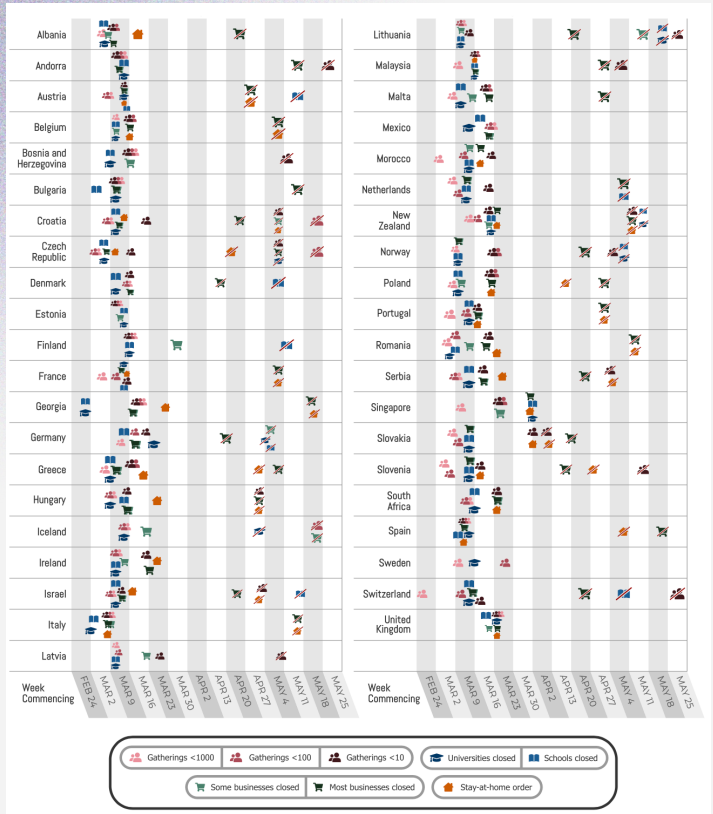
Non-pharmaceutical interventions (NPIs)

Pharmaceutical interventions

When? **How?** **What?** → Overall and long-term disease risks

1. Xia, S., Zhou, X. N., & Liu, J. (2017). Systems thinking in combating infectious diseases. *Infectious diseases of poverty*, *6*(05), 57-63.
2. Ren, J., Liu, M., Liu, Y., & Liu, J. (2023). TransCode: Uncovering COVID-19 transmission patterns via deep learning. *Infectious Diseases of Poverty*, *12*(1), 14.

Practical Challenges



To effectively respond to **rapidly evolving epidemic dynamics**

- ❌ Manually designing control strategies solely based on expert knowledge
- ❌ Exhaustively selecting intervention policies through the traditional simulation method

When?

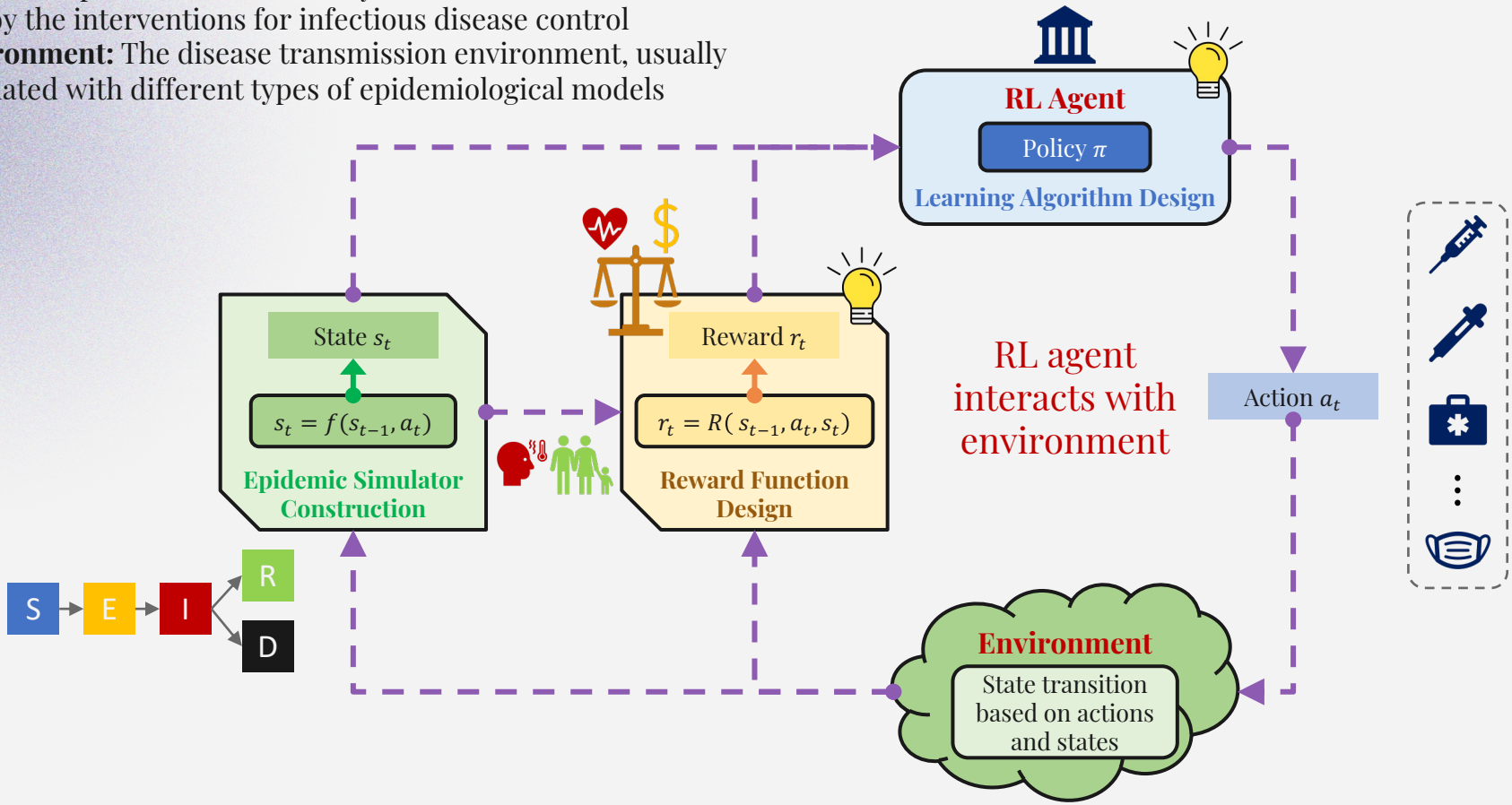
How?

What?

⇨ **Sequential decision-making problem** ⇔ **Reinforcement Learning (RL)**

1. Brauner, J. M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, J., Gavenčiak, T., ... & Kulveit, J. (2021). Inferring the effectiveness of government interventions against COVID-19. *Science*, *371*(6531), eabd9338.

- **Agent:** The public health authority or related sector to issue and deploy the interventions for infectious disease control
- **Environment:** The disease transmission environment, usually simulated with different types of epidemiological models



- **States:** Disease transmission states (e.g., susceptible, exposed, infected, recovered, and death)
- **Reward:** Defined objectives of disease control (such as decreases in infection and death or a negative economic loss)
- **Action:** Defined the intervention types and intensity

02

Literature Search Methodology

Review Scope and Methodology

- **Search sources/databases**: PubMed
- **Search temporal range**: 2020 to July 8, 2025
- **Search keywords**:
  - ("reinforcement learning") AND (("infectious disease" control) OR (epidemic control) OR ("infectious disease" intervention strategy))
  - Search keywords in full paper: About 49 results
- After being screened by the title and abstract and checking accessibility of the full text: **19**

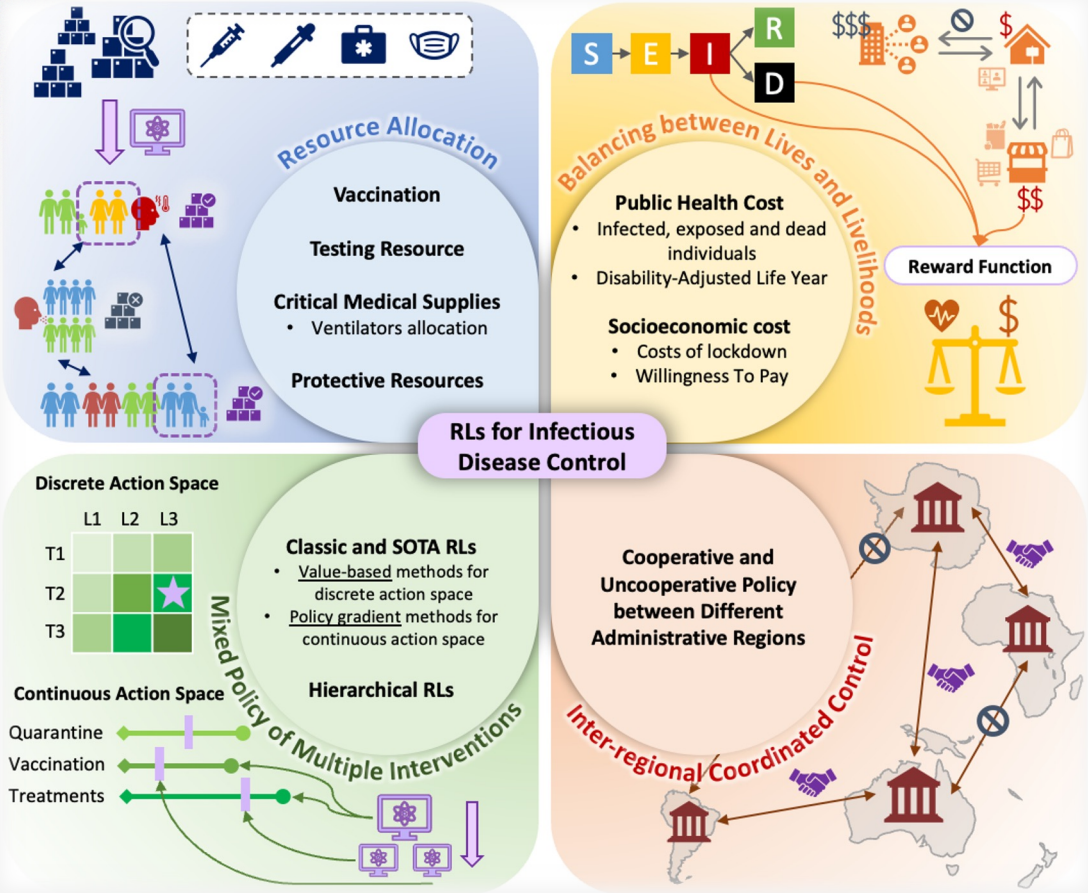
03

Survey Taxonomy

Taxonomy Framework

Resource Allocation

Mixed Policy of Multiple Interventions

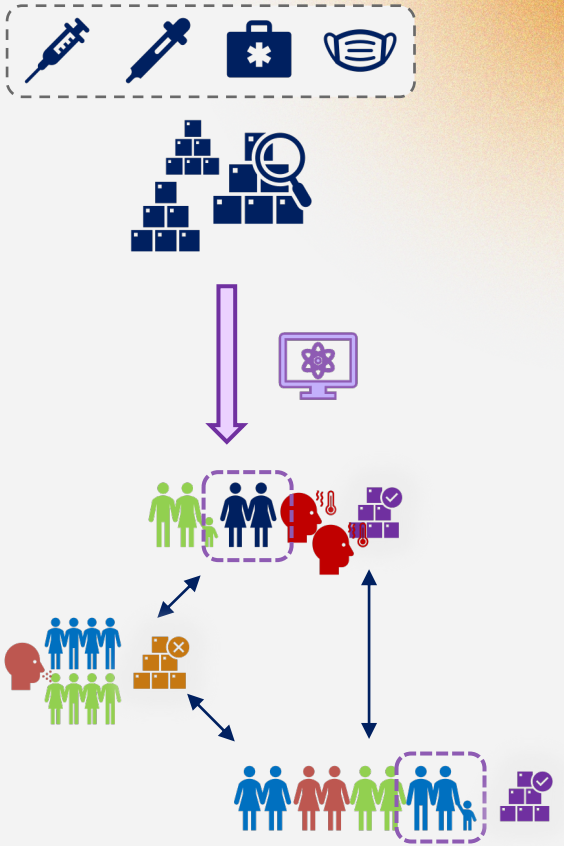


Balancing between Lives and Livelihoods

Inter-regional Coordinated Control

Resource Allocation

- **Key:** Allocating **limited resources** to high-risk or high-priority population or regions
- The RL algorithms learn their policy with the **constraint of given budgets**
- **Formulating the problem** with various requirements:
  - **Node (group) identification problem in the complex network** with selecting certain number of individuals
  - **Multi-armed bandit** for testing certain types of travelers at the border to maximize the testing efficiency
  - …



Resource Allocation

- **Limited resources:**
  - Vaccination
  - Testing resource
  - Critical medical supplies
  - Protective resources

TABLE I
SUMMARY OF THE TYPES OF LIMITED RESOURCE AND UTILIZED RL METHODS IN THE RESOURCE ALLOCATION.

Refer.	Limited Resource Allocation	RL Methods
[11] 2024	Vaccinating 20% individuals at social network	Q-learning
[12] 2024	Distributing vaccines to different census block zones, given a fixed number of vaccinations	Actor-critic
[13] 2021	Conducting PCR testing for travelers at 40 entry points on Greek borders with 400 health workers, 32 laboratories across the country	Multi-armed bandit
[14] 2021	Exchanging vacant ventilators based on the total number of initial ventilators of all states	Value iteration and Q-learning
[15] 2020	Distributing protective resources to a set of individuals at the social network	Q-learning

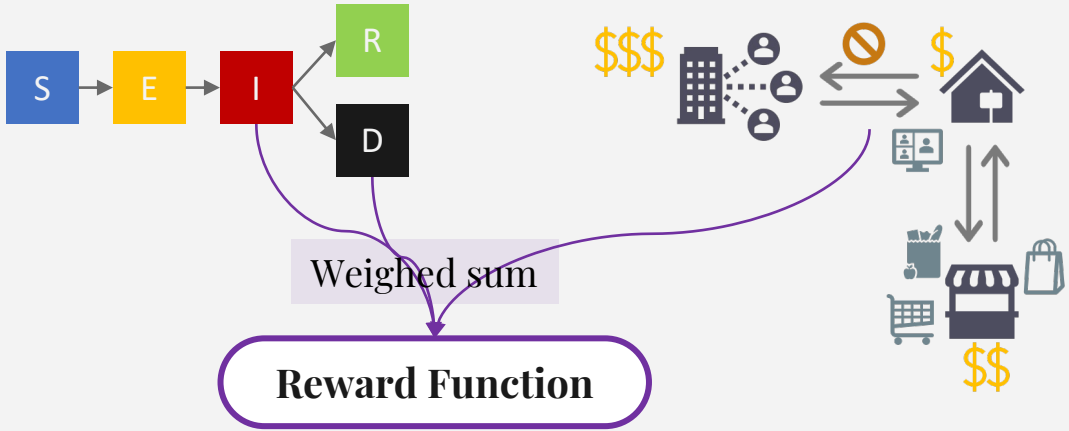
Balancing between Lives and Livelihoods

- **Key:** Learning the policy balancing between **multiple objectives**
- **Designing reward functions** to consider effects of multiple factors



Loss of health
Infected, exposed and dead individuals
Disability-Adjusted Life Year (DALY)
…

Economic cost
Costs of lockdown
Willingness To Pay
…



Balancing between Lives and Livelihoods

TABLE II

SUMMARY OF INTERVENTION TYPES AND VALUE RANGES, HEALTH RISKS CONCERNED BY PUBLIC HEALTH, AND SOCIOECONOMIC EFFECTS CAUSED BY INTERVENTIONS, INTEGRATING TO BALANCE BETWEEN LIVES AND LIVELIHOODS. THE ARROWS DENOTE THE GOAL OF INDICATOR OPTIMIZATION, WHERE ↑ DENOTES AS HIGHER AS BETTER (MAXIMIZING) AND ↓ DENOTES AS LOWER AS BETTER (MINIMIZING).

Refer.	Interventions	Public Health Effects	Socioeconomic Effects
[16] 2020	Region lockdown (Binary 0 or 1)	Number of infected population and deaths (↓)	Total costs of each lockdown day (↓)
[17] 2020	Movement restrictions (3-level percentage {0, 25%, 75%})	Number of deaths (↓)	Ratio of economic benefits brought by mobility (↑)
[18] 2021	Multi-modal inter-city travel ban (Binary 0 or 1)	Number of confirmed cases and deaths (↓)	Total retail sales of consumer goods contributed by traffic turnover volume (↑)
[19] 2022	Allowed inter-community mobility (Percentage $[0, 100\%]$)	Number of hospitalized population (↓)	Difference between the mobility demand and the actual mobility retention (↓)
[20] 2021	Vaccination for susceptible population (Percentage $[0, 100\%]$)	Number of infected and exposed individuals (↓)	Ratio of vaccinated people (↓)
[21] 2022	Social Distancing (SD) level (Decimal within $[0, \mathrm{SD}_{max}]$)	Disability-Adjusted Life Year (DALY) losses averted by SD intervention (↑)	Total amount of willingness to pay for each averted DALY (↓)

Mixed Policy of Multiple Interventions

Key: Learning the optimal combination of

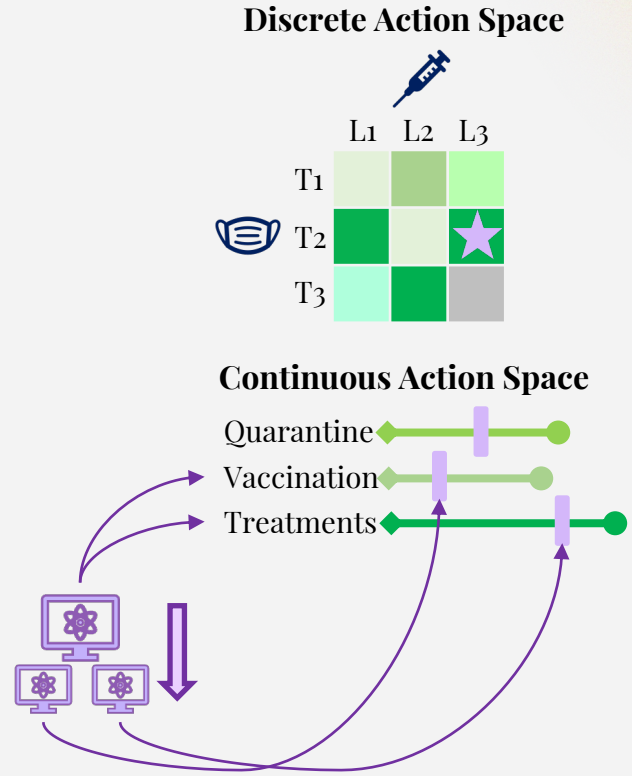
multiple interventions

RL Algorithms: Classic and SOTA RLs

- Value-based methods for discrete action space
- Policy gradient methods for continuous action space

Hierarchical RLs

- Choose the types of interventions
- Determine the intensity of them



Mixed Policy of Multiple Interventions

TABLE III
SUMMARY OF TARGETED DISEASES, MULTIPLE INTERVENTIONS, AND ADOPTED RL METHODS FOR EACH WORK IN THE MIXED POLICY OF MULTIPLE INTERVENTIONS.

Refer.	Diseases	Interventions & Action Space	RLs
[26] 2021	HIV	Discrete domain: $2 \times 3 \times 2 \times 3$ 1. Diagnostic rate for HETs (Decrement of unaware population rate): $\{0\%, -2.5\%\}$ 2. Retention-in-care rates for HETs (Increment of ART proportion): {0%, 10%, 20%} 3. Diagnostic rate for MSM: same as 1 4. Retention-in-care rates for MSM: same as 2	Q-learning
[27] 2021	COVID-19	Discrete domain: 3×3 1. Domestic lockdown: {L0, L1, L2} 2. Travel restrictions: {T0, T1, T2}	D3QN
[28] 2024	COVID-19	Continuous domain for each: $[0, 1]$ 1. Quarantine of susceptible 2. Vaccination 3. Treatments	DDPG
[29] 2024	COVID-19	Continuous domain: 1. Moving distances: $[1, 5]$ 2. Radius of screening area: $[0, 10]$	A3C
[30] 2022	COVID-19	Continuous domain for each: $[-1, 1]$ 1. Cancel public events 2. Close public transport 3. International travel controls 4. Public information campaigns 5. Gathering restrictions 6. Internal movement restrictions 7. School closing 8. Stay at home requirements 9. Workplace closing	DDPG
[31] 2023	COVID-19	Discrete domain: 9 actions 1. Do nothing 2. Testing, contact tracing, and quarantine 3. Close schools and non-essential workplaces 4. Mandatory mask 5. Testing, contact tracing, quarantine, close schools, and non-essential workplaces 6. testing, contact tracing, quarantine, and mandatory mask 7. Vaccination 8. Vaccination and mandatory mask 9. Total lockdown	DRL
[32] 2023	COVID-19	Continuous and discrete domain: 1. Mobility constraint: $[0, 100\%]$ 2. Temporary medical resources: {L, M, H} 3. Necessities supply: {Non,L,M}	Hierarchical PPO

Inter-regional Coordinated Control

- **Key:** Jointly optimizing the policies of multiple regions for inter-regional coordination
- **Limited works to research this topic!**
- Explore cooperative and the Uncooperative Policy between Different Administrative Regions → Multi-agent Reinforcement Learning

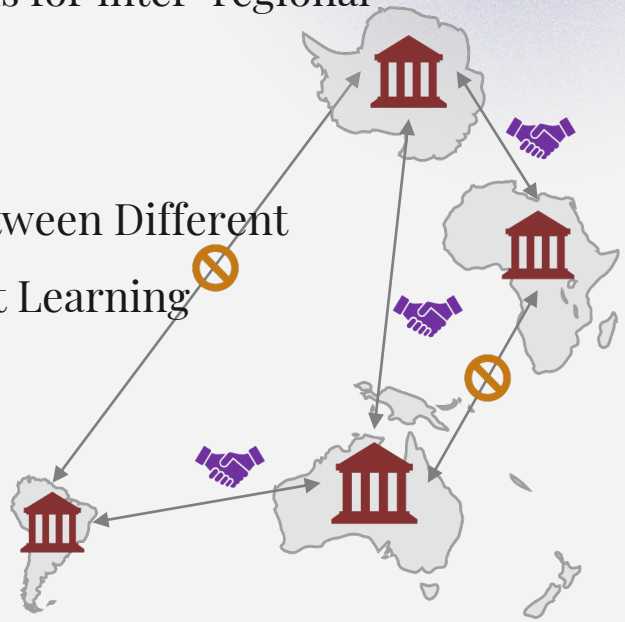


Table 2. Summary of the scenarios studied.

Scenario	Objective function	Number of jurisdictions	Policy	Travel from B to A	Initiation of optimal policy (days)
Scenario 1	2-term	Single jurisdiction, A	A optimal policy	Not applicable	30, 60, 75, 90, 95, 100, 105,110
Scenario 2	2-term	Two jurisdictions, A and B	A optimal policy, B no intervention	5%	30, 60, …, 110
Scenario 3	2-term	Two jurisdictions	A optimal policy, B no intervention	10%	30, 60, …, 110
Scenario 4	2-term	Two jurisdictions	A optimal policy, B optimal policy	5%	30, 60, …, 110
Scenario 5	2-term	Two jurisdictions	A optimal policy, B optimal policy	10%	30, 60, …, 110
Scenario 6	3-term	Single jurisdiction	A optimal policy	Not applicable	30, 60, …, 110
Scenario 7	3-term	Two jurisdictions	A optimal policy, B no intervention	10%	30, 60, …, 110

1. Khatami, S. N., & Gopalappa, C. (2023). Deep reinforcement learning framework for controlling infectious disease outbreaks in the context of multi-jurisdictions. *Mathematical Biosciences and Engineering*, 20(8).

04

Conclusions and Future Directions

Conclusions and Future Directions

Limitations of our survey:

- Did not cover all of the relevant literature as we defined the certain search scope and keywords

Promising future directions:

- Effectively learning **intervention combinations** in the large solution space
- Enhancing **coordination between multiple administrative regions**
- Constructing **standard and general benchmarks** for uniform comparison of diverse algorithms and indicators

Thanks!

Do you have any questions?

csmtliu@comp.hkbu.edu.hk

CREDITS: This presentation template was created by **Slidesgo**, and includes icons by **Flaticon**, and infographics & images by **Freepik**