Assignment: Agent-Based SIR Simulation with Intervention Optimization via RL

Course: State-of-the-art in AI
Instructor: Teddy Lazebnik

September 29, 2025

Abstract

In this assignment you will (i) build an *agent-based simulation* (ABS) of the classical SIR epidemic model, (ii) design at least one *intervention mechanism* that modifies transmission dynamics, and (iii) implement a simple *reinforcement learning (RL)* agent that learns an intervention policy to optimize a defined objective. Your submission must include three plots: (1) SIR dynamics produced by your ABS, (2) a scatter plot relating the transmission parameter β to the basic reproduction number $\mathcal{R}_0$, and (3) an RL training curve. This document provides a step-by-step guide, deliverables, and a rubric.

Contents

1 Background **2**
1.1 SIR (ODE) Reference 2
1.2 Agent-Based SIR Intuition 2

2 Part A (30%): Build an Agent-Based SIR **2**

3 Part B (25%): Design an Intervention **3**

4 Part C (30%): RL for Intervention Optimization **3**

5 Part D (15%): Experiments & Required Plots **4**

6 Reproducibility Checklist **5**

7 What to Submit **5**

8 Suggested Parameter Defaults (Optional) **5**

---

1 Background

1.1 SIR (ODE) Reference

The classical SIR model partitions a closed population of size N into Susceptible $S(t)$, Infectious $I(t)$, and Removed/Recovered $R(t)$:

$$\frac{dS}{dt} = -\beta \frac{SI}{N}, \tag{1}$$

$$\frac{dI}{dt} = \beta \frac{SI}{N} - \gamma I, \tag{2}$$

$$\frac{dR}{dt} = \gamma I. \tag{3}$$

Here β is the transmission rate and γ is the recovery rate. In a well-mixed homogeneous population, $\mathcal{R}_0 = \beta/\gamma$.

1.2 Agent-Based SIR Intuition

In an ABS, we simulate individuals (agents) who interact according to a contact process. Each infectious agent attempts infectious contacts with susceptible agents; successful transmission changes the target agent to infectious. Infectious agents recover stochastically.

Discrete-Time Mapping. With a time step Δt, one can model per-contact transmission probability $p_{\text{trans}} = 1 - e^{-\beta\,\Delta t}$ and per-step recovery probability $p_{\text{rec}} = 1 - e^{-\gamma\,\Delta t}$. If each infectious agent makes C effective contacts per step (or contacts are derived from a network), the ABS can approximate the ODE model when N is large and mixing is near-homogeneous.

2 Part A (30%): Build an Agent-Based SIR

Step A1: Model Assumptions

Choose *one* of the following contact structures and state it clearly in your report:

a) **Well-mixed**: each infectious agent selects C contacts uniformly at random from the population (without replacement) each step.

b) **Spatial grid**: agents occupy cells; contacts occur with Moore/ von Neumann neighbors.

c) **Network**: contacts follow a static or dynamic graph (e.g., Erdős–Rényi or small-world).

Step A2: Parameters

- Population size: N (e.g., $N \in \{1000, 5000\}$).
- Time step: Δt (e.g., $\Delta t = 1$ day or $1/24$ day (i.e., hour)).
- Transmission rate: β (grid for experiments below).
- Recovery rate: γ (e.g., $\gamma = 1/7\,\text{day}^{-1}$).
- Contacts per step (if applicable): C (e.g., $C \in \{5, 10\}$).

Step A3: State Update (Pseudocode)

```
for each agent i:
  if state[i] == Infectious:
    # infectious contacts
    contacts = sample_without_replacement(others, C)
    for j in contacts:
      if state[j] == Susceptible:
        if Bernoulli(1 - exp(-beta * delta_t)):
          next_state[j] = Infectious
    # recovery
    if Bernoulli(1 - exp(-gamma * delta_t)):
      next_state[i] = Recovered
# apply all changes simultaneously
state <- next_state
```

Listing 1: One time-step update for well-mixed ABS

You may vectorize or otherwise optimize; update synchronously each step.

Step A4: Sanity Checks

- With small $I(0)$ and no intervention, early growth should be approximately exponential.
- Aggregate your ABS counts (S, I, R) and compare qualitatively with ODE SIR using the same (β, γ).

3 Part B (25%): Design an Intervention

Design *at least one* intervention mechanism that influences transmission. Examples:

- **Contact reduction** $u_t \in [0, 1]$: replace C with $(1 - u_t)\, C$.
- **Transmission reduction**: effective per-contact probability becomes $(1 - u_t)\, p_{\text{trans}}$ (e.g., masking).
- **Vaccination**: move a fraction v_t of susceptibles to Recovered at specified times.
- **Targeted isolation**: detected infectious agents make zero contacts for d days.

Define a **cost** for interventions. For example, per-step cost

$$\text{cost}_t = \lambda_{\text{epi}} \cdot \text{newInfections}_t \; + \; \lambda_{\text{soc}} \cdot u_t^2, \tag{4}$$

with tunable weights $\lambda_{\text{epi}}, \lambda_{\text{soc}} > 0$.

4 Part C (30%): RL for Intervention Optimization

Step C1: MDP Specification

Define an MDP where:

- **State** s_t: a compact summary such as $\big(S_t/N,\, I_t/N,\, R_t/N,\, t/T\big)$ and optionally previous action u_{t-1}.
- **Action** a_t: choose intervention intensity u_t from a discrete set, e.g., $\{0, 0.25, 0.5, 0.75, 1.0\}$, or a small continuous range.

- **Transition**: given by your ABS with intervention applied.
- **Reward**: $r_t = -\text{cost}_t$ (i.e., minimize epidemiological burden and intervention cost).
- **Episode length**: e.g., $T = 100$–200 steps.

Step C2: Algorithm Choice

Use any *simple* RL method appropriate to your action space:

- Discrete u_t: tabular Q-learning or SARSA with ϵ-greedy exploration.
- Low-dimensional continuous u_t: policy gradient with a tiny network or discretize actions.

Document hyperparameters (learning rate, discount factor γ_{RL}, exploration schedule, episode count). Train until returns stabilize.

Step C3: Baselines

Compare the learned policy to a fixed baseline (e.g., no intervention, or constant $u_t = u_0$). Report total cost and final outbreak size.

5 Part D (15%): Experiments & Required Plots

All plots must have labeled axes, units, legends, and informative captions.

Plot 1: SIR Dynamics from ABS

Run the ABS with chosen (β, γ) and no intervention. Plot $S(t)$, $I(t)$, $R(t)$ vs. time.

Plot 2: Scatter of β vs. $\mathcal{R}_0$

Protocol to estimate $\mathcal{R}_0$ empirically: For each β in a grid (e.g., $\{0.05, 0.075, 0.10, \ldots, 0.30\}$ day^{-1}):

i) Initialize with one infectious index case and $N - 1$ susceptibles.

ii) Simulate until the index case recovers; record the *number of secondary infections directly caused by that index*.

iii) Repeat for R independent runs (e.g., $R = 100$) and average to obtain $\widehat{\mathcal{R}_0}(\beta)$.

Plot the scatter of $(\beta, \widehat{\mathcal{R}_0})$ and, optionally, a regression line. Also compare to the theoretical curve $\mathcal{R}_0 = \beta/\gamma$ when appropriate.

Plot 3: RL Training Curve

Plot the episode return (sum of rewards) vs. training episode, with a moving average (e.g., window 20) to visualize learning progress. Optionally add baseline return as a horizontal reference.

6 Reproducibility Checklist

- Fixed random seeds; specify language and versions (e.g., Python, packages).
- Document all parameters: N, Δt, β, γ, C, contact structure, RL hyperparameters.
- Number of runs per configuration and aggregation method (mean $\pm$ 95% CI, etc.).
- Provide a script/notebook that regenerates all figures end-to-end.

7 What to Submit

A PDF report including:

A. ABS implementation details and sanity checks - up to a page.

B. Intervention design and cost definition - up to half a page.

C. RL formulation - up to half a page.

D. Results: the three required plots.

8 Suggested Parameter Defaults (Optional)

As a starting point (modify as needed): $N = 5000$, $\Delta t = 1$ day, $\gamma = 1/7$ day^{-1}, $C = 8$. For Plot 2, sweep β in $[0.05, 0.30]$ day^{-1} with step 0.025. For RL, episodes $T = 150$, discount $\gamma_{\text{RL}} = 0.99$, learning rate $\alpha = 0.1$, ϵ-greedy starting at 0.2 and decaying to 0.01.