

Reinforcement Learning in Health

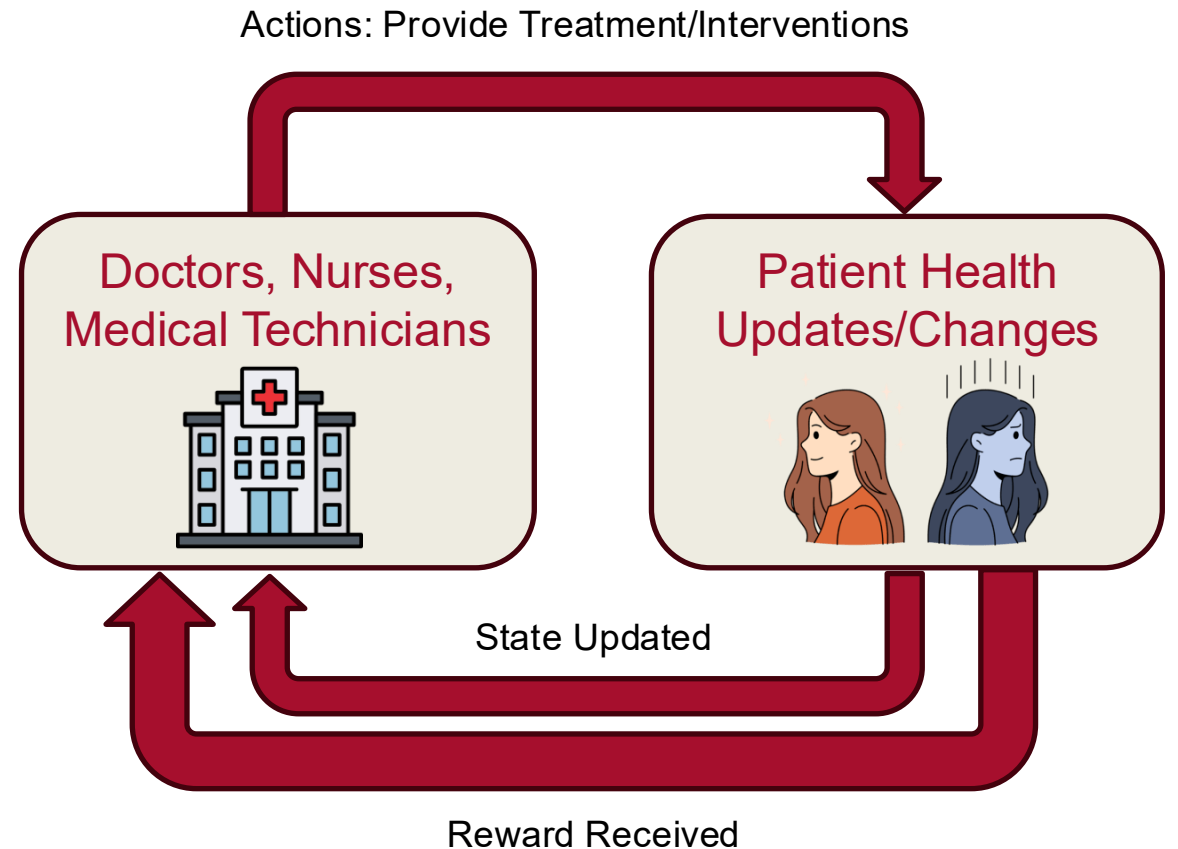
Saman Khamesian, Hassan Ghasemzadeh

College of Health Solutions
Embedded Machine Intelligence Lab
Arizona State University



Reinforcement Learning in Health

- So far, we've seen how an agent can learn to navigate a small game world.
- But real life is not a game (especially when it comes to healthcare).
- Can these same ideas help patients manage chronic conditions or help doctors make better decisions?



Reinforcement Learning in Health

- Healthcare is a **high-stakes domain** where we can't explore freely like in games.
- **Key Challenges:**
 -  Ethical constraints
 -  Delayed rewards (e.g., long-term health effects)
 -  Limited exploration
 -  Sparse, noisy or biased data
 -  Safety (we can't experiment on patients freely)
- That's why reinforcement learning must rely on **simulations** and **careful design**.

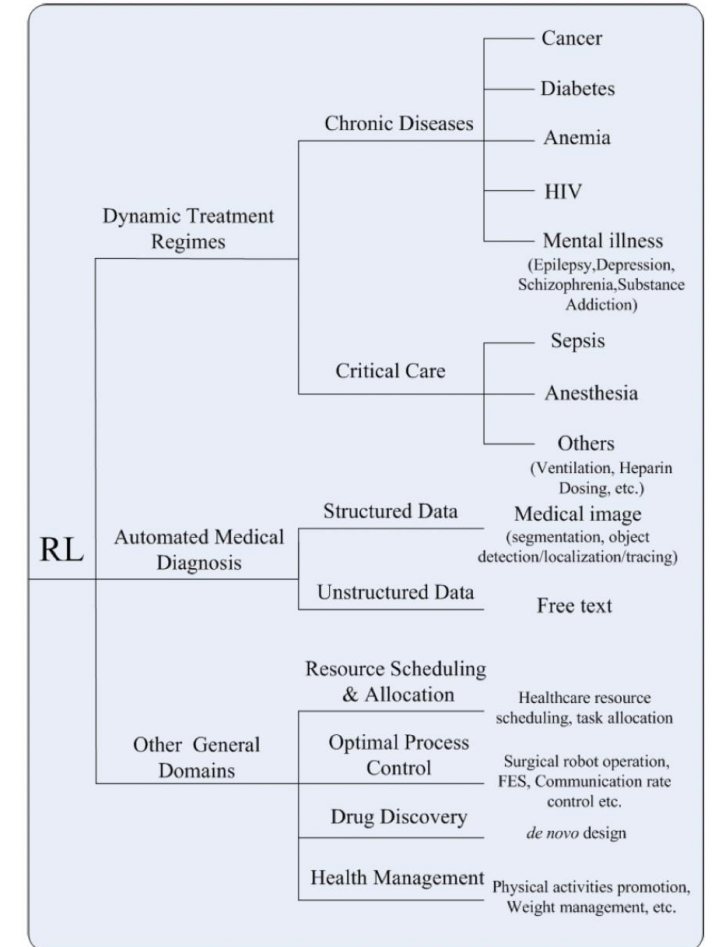
Reinforcement Learning in Health

 **Diabetes Management:** RL agents adjusting insulin in simulated environments to mimic pancreas function.

 **Sepsis Treatment:** Deep RL suggesting drug doses for patients in ICU (e.g., from MIT MIMIC dataset).

 **Prosthetic Control:** Agents helping optimize movement and gait in prosthetic limbs.

 **Drug Discovery:** RL used in optimizing molecule design for new treatments.

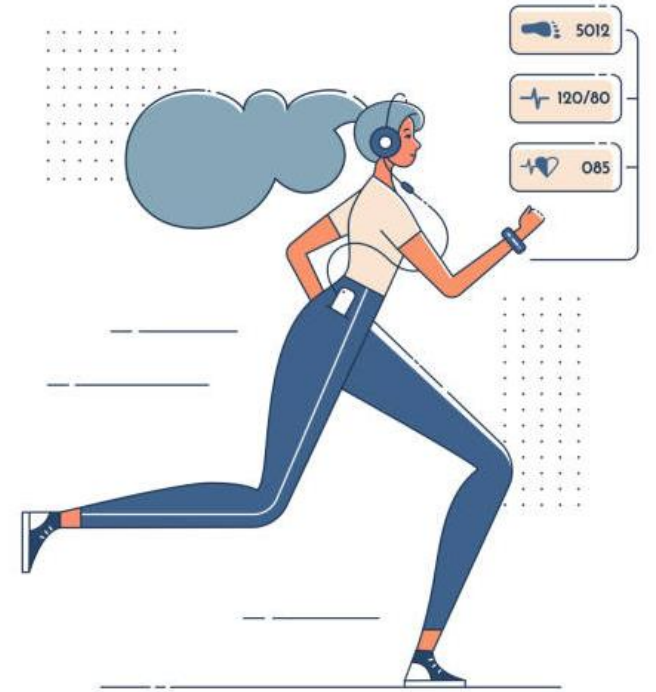


Example for Physical Activity

- Imagine we're building a reinforcement learning agent to encourage people to walk more during the day.
- The agent runs on a smartphone. It can track:
 -  Step count
 -  Time of day
 -  GPS location
 -  User's walking history
- It learns when to suggest short walks based on daily patterns and habits.

Environment

- Now that we've talked about encouraging walking, let's take a closer look at the environment the agent interacts with.
- In this case, the **environment is the real world around the user** — their daily routine, movements, and context.
- The agent's goal is to help them walk more by learning when and how to suggest activity.
- It acts like a **smartphone-based health coach** — running quietly during the day, observing patterns, and nudging the user a few times when it sees the opportunity.



State Space

- The agent observes a combination of:

Feature	Description
steps_today	Number of steps so far today
time_of_day	Encoded as morning, afternoon, evening
location_type	Home, work, park, unknown (based on GPS cluster)
last_action_taken	What the agent did last (to avoid repeat prompts)
previous_day_steps	Total steps yesterday
current_weather	Sunny, rainy, hot (from weather API)
response_to_last_prompt	Whether the user walked after previous suggestion (binary)

Action Space

- The RL agent chooses **at most 3–4 times per day** from the following:

Action	Description
SEND_PROMPT_NOW	Notify user: “Would you like to go for a short walk now?”
SUGGEST_LATER	Suggest scheduling a walk later today (e.g., “How about a 10-minute walk this evening?”)
SHOW_PROGRESS	Display stats: “You’re only 1,500 steps away from today’s goal!”
DO_NOTHING	Skip intervention — helps avoid over-prompting

Reward Function

- A simple reward design combining short-term behavior and long-term goals:

$$R_t = \Delta\text{steps}_t + 5 \times 1[\text{goal_achieved}_t] - 2 \times 1[\text{prompt_ignored}_t]$$

- **Explanation:**
 - Δsteps_t : Additional steps within 2 hours of the action
 - goal_achieved : +5 if total steps exceed 7,000 (or target set for the user)
 - prompt_ignored : -2 penalty if user consistently ignores prompts

Questions

- **Question 1:** What would you do if your RL agent kept making the “right” suggestions — but users kept ignoring them?
- **Question 2:** How would you redesign the *DO_NOTHING* action to be smarter?



Hassan Ghasemzadeh

Director
Arizona State University



Saman Khamesian

Graduate Research Associate
Arizona State University

Embedded Machine Learning Lab (EMIL)

Hassan.Ghasemzadeh@asu.edu

ghasemzadeh.com



Thank You!