StepCountJITAI: simulation environment for RL with application to physical activity adaptive intervention

Karine and Benjamin M. Marlin

University of Massachusetts Amherst

Motivations and challenges

- There is an increasing interest in using Reinforcement learning (RL) to learn policies for just-in-time adaptive interventions (JITAIs).
- Physical activity adaptive intervention can be framed as **RL**: a mobile health app (**agent**) sends messages (**actions**) to a participant (**environment**) to encourage the participant to exercise more (**reward** = walking step count).

However, there are challenges:

- Cost and time constraints of real studies → limited training data.
- Lack of simulation environments for physical activity adaptive intervention.

Contributions

1. StepCountJITAI is a simulation environment that models stochastic behavioral dynamics and context uncertainty, with parameters to control stochasticity.
2. StepCountJITAI can be used to develop new RL algorithms for physical activity adaptive intervention and accelerate research work.
3. StepCountJITAI is open-source and uses standard API for RL (i.e., gymnasium). Code and quickstart are available at: github.com/reml-lab/StepCountJITAI

StepCountJITAI simulation environment actions, states, parameters

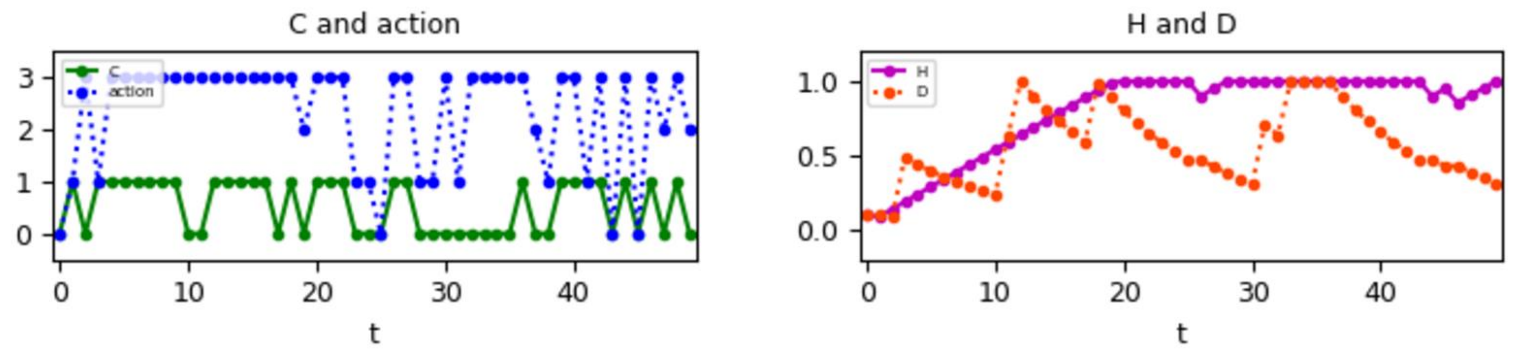
Action value	Description
$a = 0$	No message is sent to the participant.
$a = 1$	A non-contextualized message is sent to the participant.
$a = 2$	A message customized to context 0 is sent to the participant.
$a = 3$	A message customized to context 1 is sent to the participant.

Variable	Description	Values
c_t	True context.	$\{0, 1\}$
p_t	Probability of context 1.	$[0, 1]$
l_t	Inferred context.	$\{0,1\}$
d_t	Disengagement risk level.	$[0, 1]$
h_t	Habituation level.	$[0, 1]$
s_t	Step count.	$\mathbb{R}^+$

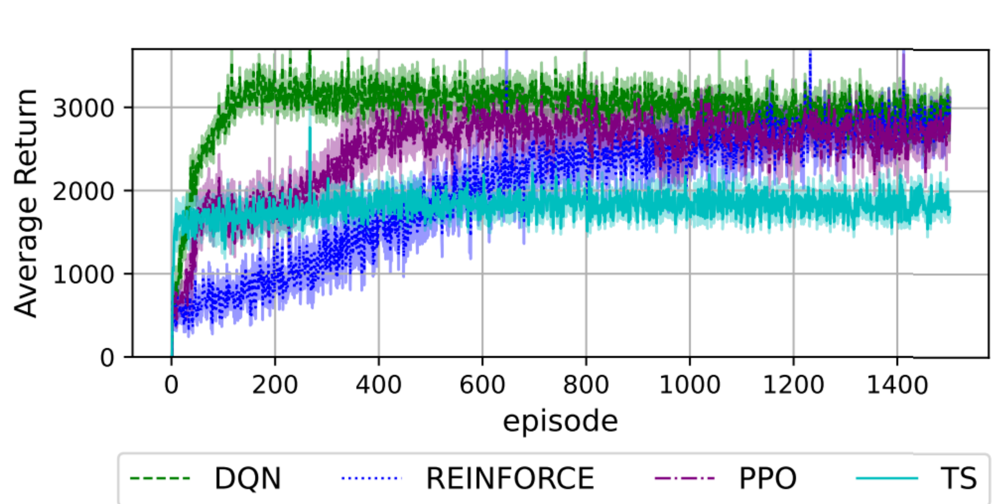
Parameter	Description
σ	Context uncertainty. The default value is $\sigma = 0.4$
δ_d	Disengagement risk decay. The default value is $\delta_d = 0.1$.
δ_h	Habituation decay. The default value is $\delta_h = 0.1$.
ϵ_d	Disengagement risk increment. The default value is $\epsilon_d = 0.4$.
ϵ_h	Habituation increment. The default value is $\epsilon_h = 0.05$.
ρ_1	$a_t = 1$ base step count. The default value is $\rho_1 = 50$.
ρ_2	$a_t = c_t + 2$ base step count. The default value $\rho_2 = 200$.
m_s	Base step count. The default value is $m_s = 0.1$.

Parameter	Description
σ_s	Parameter to control the spread of the Gamma distribution for s_t.
a_{hd}	Parameter to control the spread of the Uniform distributions for h_t and d_t.
a_{de}	Parameter to control the spread of the Uniform distributions for $\delta_d, \epsilon_d, \delta_h$ and ϵ_h.

Experiments

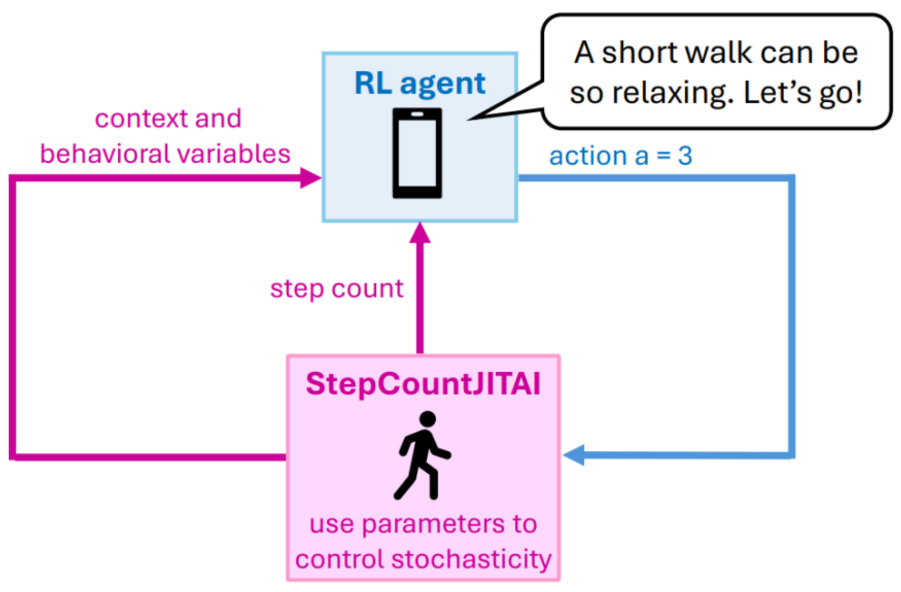


Example of traces for action, context C and behavioral dynamics: habituation level H and disengagement risk D, (traces are obtained by running a standard Thompson Sampling).



Example of average returns using StepCountJITAI with RL methods.

Methods



Overview of StepCountJITAI in an RL loop.
StepCountJITAI is a simulation environment that models stochastic behavioral dynamics and context uncertainty, with parameters to control stochasticity.

StepCountJITAI simulation environment behavioral dynamics

Stochasticity is added in the behavioral dynamics (stochastic equations are in blue). The spreads of the distributions are controlled by the parameters (e.g., a_{de}, a_{hd}, σ_s).

$$\delta_h \sim Uniform\big((1 - \frac{a_{de}}{2})\hat{\delta_h}, (1 + \frac{a_{de}}{2})\hat{\delta_h}\big) \qquad \delta_d \sim Uniform\big((1 - \frac{a_{de}}{2})\hat{\delta_d}, (1 + \frac{a_{de}}{2})\hat{\delta_d}\big)$$

$$\epsilon_h \sim Uniform\big((1 - \frac{a_{de}}{2})\hat{\epsilon_h}, (1 + \frac{a_{de}}{2})\hat{\epsilon_h}\big) \qquad \epsilon_d \sim Uniform\big((1 - \frac{a_{de}}{2})\hat{\epsilon_d}, (1 + \frac{a_{de}}{2})\hat{\epsilon_d}\big)$$

$$c_{t+1} \sim Bernoulli(0.5), \quad x_{t+1} \sim \mathcal{N}(c_{t+1}, \sigma^2)$$

$$p_{t+1} = P(C = 1|x_{t+1}), \quad l_{t+1} = p_{t+1} > 0.5$$

$$\hat{d}_{t+1} = \begin{cases} d_t & \text{if } a_t = 0 \\ (1 - \delta_d) \cdot d_t & \text{if } a_t \in \{1, c_t + 2\} \\ \min(1, d_t + \epsilon_d) & \text{otherwise} \end{cases}$$

$$d_{t+1} \sim Uniform\big((1 - \frac{a_{hd}}{2})\hat{d}_{t+1}, (1 + \frac{a_{hd}}{2})\hat{d}_{t+1}\big)$$

$$\hat{h}_{t+1} = \begin{cases} (1 - \delta_h) \cdot h_t & \text{if } a_t = 0 \\ \min(1, h_t + \epsilon_h) & \text{otherwise} \end{cases}$$

$$h_{t+1} \sim Uniform\big((1 - \frac{a_{hd}}{2})\hat{h}_{t+1}, (1 + \frac{a_{hd}}{2})\hat{h}_{t+1}\big)$$

$$\hat{s}_{t+1} = \begin{cases} m_s + (1 - h_{t+1}) \cdot \rho_1 & \text{if } a_t = 1 \\ m_s + (1 - h_{t+1}) \cdot \rho_2 & \text{if } a_t = c_t + 2 \\ m_s & \text{otherwise} \end{cases}$$

$$s_{t+1} \sim Gamma\big((\frac{\hat{s}_{t+1}}{\sigma_s})^2, \frac{\sigma_s^2}{\hat{s}_{t+1}}\big)$$

For the stochastic version using Beta distribution, please refer to our paper.

Conclusion

✓ StepCountJITAI is a simulation environment that models stochastic behavioral dynamics and context uncertainty, with parameters to control stochasticity.

✓ StepCountJITAI can be used to create new RL algorithms for physical activity adaptive intervention and accelerate research work.

✓ StepCountJITAI offers a solution to the lack of simulation environments for physical activity adaptive intervention.

References

- Karine Karine and Benjamin Marlin (2024). "StepCountJITAI: simulation environment for RL with application to physical activity adaptive intervention". In: NeurIPS 2024 Workshop on Behavioral Machine Learning.
- Karine Karine, Susan Murphy, Benjamin Marlin (2024). "BOTS: Batch Bayesian Optimization of Extended Thompson Sampling for Severely Episode-Limited RL Settings". In: NeurIPS 2024 Workshop on Bayesian Decision-making and Uncertainty.

This work is supported by National Institutes of Health National Cancer Institute, Office of Behavior and Social Sciences, and National Institute of Biomedical Imaging and Bioengineering through grants U01CA229445 and 1P41EB028242.