Discovering User Types: Mapping User Traits by Task-Specific Behaviors

Lars L. Ankile, Brian S. Ham, Kevin Mao, Eura Shin, Siddharth Swaroop, Finale Doshi-Velez, Weiwei Pan

Motivation

mHealth Applications Mobile health (mHealth) applications are gaining popularity as a cost-effective health intervention method.

Improvements by Personalization Users have intrinsic barriers hindering them from achieving desired outcomes, and we want to learn specific barriers using Reinforcement Learning (RL) methods to personalize.

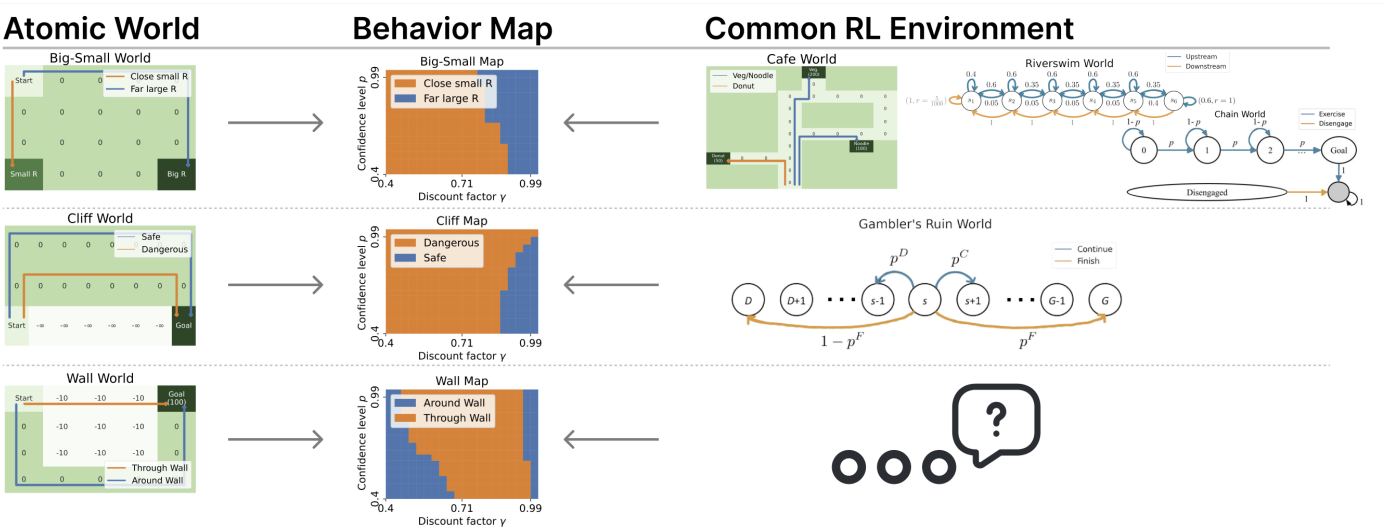
Main Contributions

Contribution 1: A tool for visualizing the relationship between user traits and the corresponding user behavior (**behavior maps**).

Contribution 2: Classify common RL worlds into small set of equivalence classes.

Contribution 3: The equivalence relation enables transferring intervention knowledge between worlds.

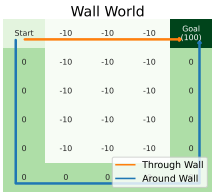
Many RL Worlds Map to Small Number of Equivalence Classes



Our Model of mHealth

World Model: an MDP $\mathcal{M} = \langle \mathcal{S}, \mathcal{A}, T, R, \gamma \rangle$

User Model: Intrinsic barriers modeled by varying $T_p^{\text{user}}, \gamma^{\text{user}}$—the user's *confidence* and *myopia*.



Confidence Parameter: Confidence, or self-efficacy, is modeled by the transition dynamics T_p^{user}, with $p \in [0, 1]$.

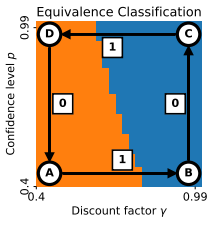
Myopia Parameter: An agent's planning ability is modeled with discount factor $\gamma^{\text{user}} \in [0, 1]$.

Impairments have strong grounding in behavioral science.

Behavior Maps and Equivalence Classes

Behavior Map: a visualization of a function from user parameters to user behavior (classes of policies)

Equivalence Class: The vector of number of behavior changes along each of the edges of the behavior map—*e.g., here:* $[1, 0, 1, 0]$.



Equivalence Allows Intervention Transfer

Many seemingly different worlds belong to the same equivalence class.

Knowledge about user behavior can transfer between worlds belonging to the same equivalence class.

Insights from behavior maps help warm-start intervention design by identifying impactful treatments and estimating effect sizes.

Future Work

- The number of equivalence classes is unknown.
- More complex worlds seem decomposable to multiple atomic worlds.
- These concepts have not been tested empirically.
- The equivalence relation might depend on the specific modeling choices to a non-trivial degree.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. IIS-2107391.