Investigating Human Priors for Playing Video Games

Rachit Dubey, Pulkit Agrawal, Deepak Pathak, Thomas L. Griffiths, Alexei A. Efros

Presenter: Arshdeep Sekhon
https://qdata.github.io/deep2Read

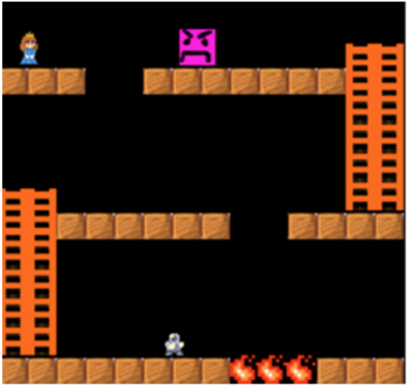
Motivation

- Humans better than computers at games because they have prior knowledge about the world
- RL requires large amounts of training data in terms of interaction with the environment
- inefficient compared to human players
- A lot of work focusses on improving sample efficiency
- This paper: Quantify role of human priors in games

Overview: Human Priors in Games

- Removing prior : Humans show difference in performance
- RL Algorithms : same performance
- RL doesn't carry any prior information about the games
- Hence, humans use prior knowledge to solve most tasks
- Ablation study on a game enviroment to mask different types of priors
- SPecific priors as well as general priors

Game



(a) Original Game

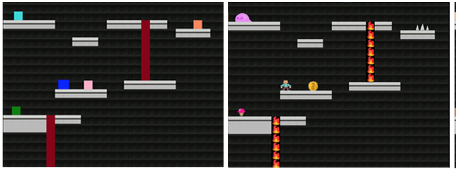
- End goal and reward remains the same
- different versions of the video game with the same underlying structure and reward

Details

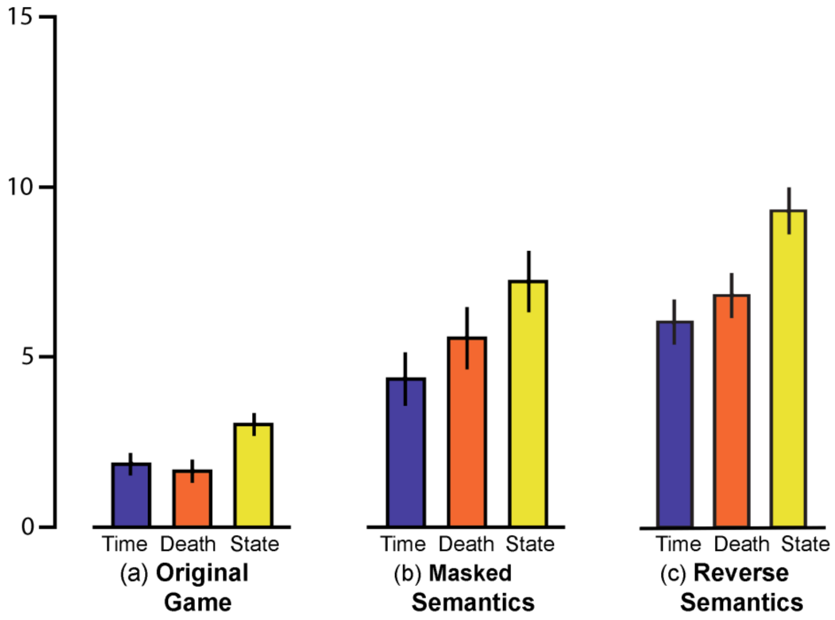
- Mechanical Turk
- no information about final goal or reward structure
- Evals:
  - total time taken to finish the game
  - total number of deaths
  - (x, y) position : states

Prior: Object Semantics

- rendered objects and ladders with blocks of uniform color
- Switch Connotations: Replace slime and spikes with ice cream etc
- enables humans to infer the latent reward structure of the game

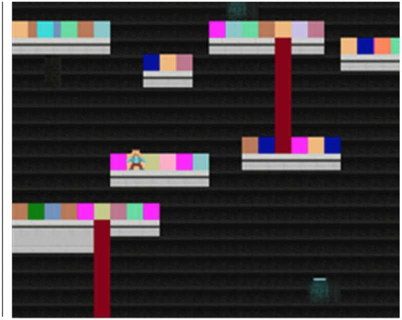


(b) Masked Semantics (c) Reverse Semantics (

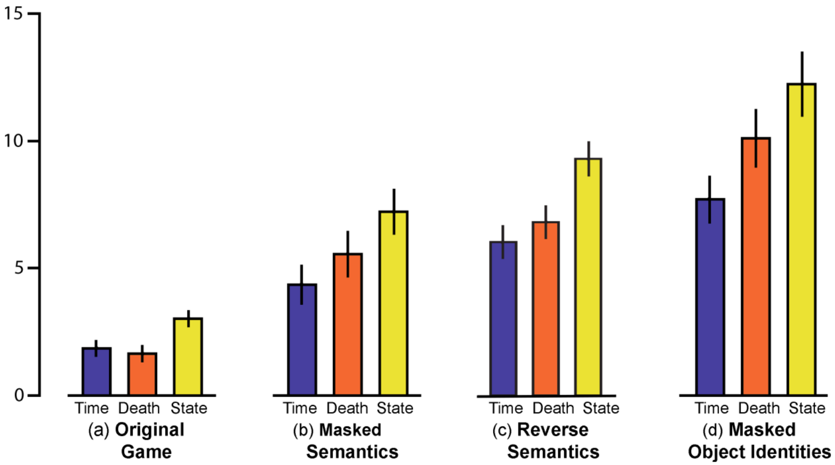

15
10
5
0
Time
Death
State
(a) **Original Game**
Time
Death
State
(b) **Masked Semantics**
Time
Death
State
(c) **Reverse Semantics**

Human Prior: Objects as sub-goals for exploration

- they are distinct from the background and seem to attract human attention
- cover each space on the platform with a block of different color to hide where the objects are

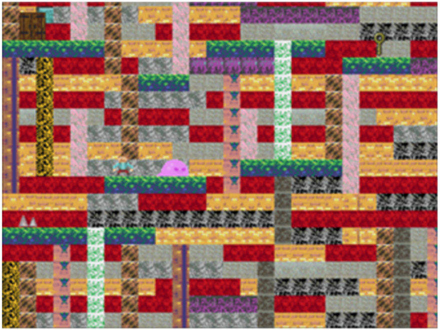


(d) Masked identity of objects

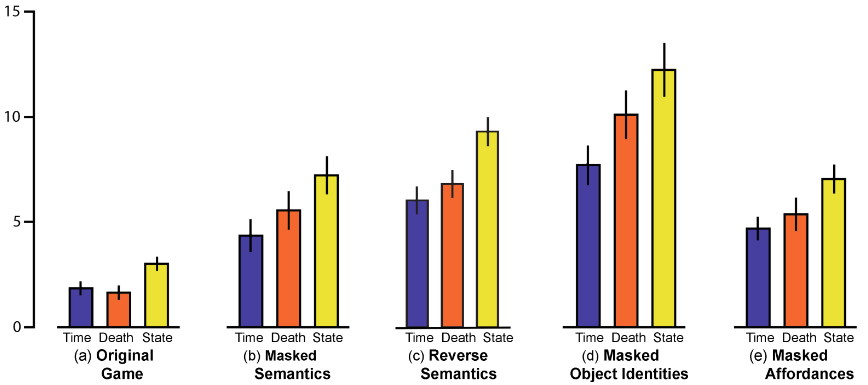

15
10
5
0
Time
Death
State
(a) Original Game
Time
Death
State
(b) Masked Semantics
Time
Death
State
(c) Reverse Semantics
Time
Death
State
(d) Masked Object Identities

General Prior: Affordances

- ladders are for climbing, can stand on platforms
- platforms and ladders afford the actions of walking and climbing
- One way to mask affordances is to fill free space with random textures, which are visually similar to textures used for rendering ladders and platforms
- objects and their semantics are clearly observable.

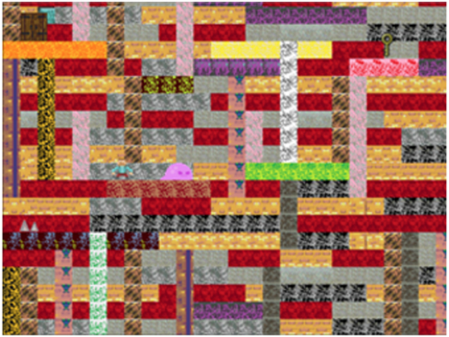


(e) Masked Affordances

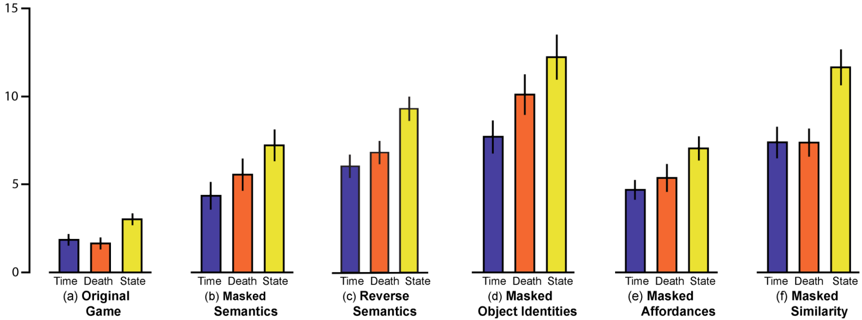

15
10
5
0
Time Death State
(a) **Original Game**
Time Death State
(b) **Masked Semantics**
Time Death State
(c) **Reverse Semantics**
Time Death State
(d) **Masked Object Identities**
Time Death State
(e) **Masked Affordances**

General Prior: Things that look similar, behave similarly

- once the player realizes that it is possible to stand on a particular texture and climb a specific texture, it is easy to use color/texture similarity to identify other platforms and ladders
- none of the platforms and ladders had the same visual texture

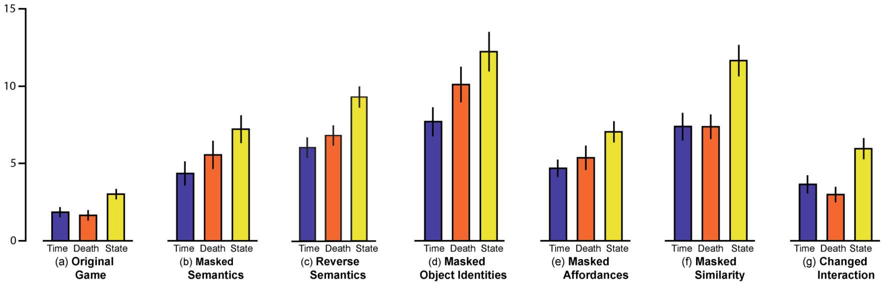


(f) Masked Visual Similarity


15
10
5
0
Time Death State
(a) Original Game
Time Death State
(b) Masked Semantics
Time Death State
(c) Reverse Semantics
Time Death State
(d) Masked Object Identities
Time Death State
(e) Masked Affordances
Time Death State
(f) Masked Similarity

Human Prior: How to interact with objects

- Knowing only what the object is
- Next, How to interact with the object
- jump over a monster, climb a ladder
- zigzag ladders


15
10
5
0
Time Death State
(a) Original Game
Time Death State
(b) Masked Semantics
Time Death State
(c) Reverse Semantics
Time Death State
(d) Masked Object Identities
Time Death State
(e) Masked Affordances
Time Death State
(f) Masked Similarity
Time Death State
(g) Changed Interaction

Ablation Study

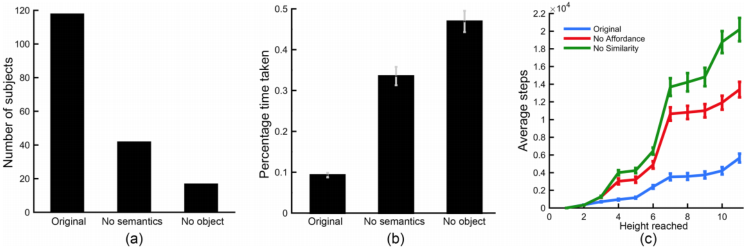


Figure: Average number of steps taken by participants to reach various vertical levels in original version, game without affordance, and game without similarity

Human Prior: Physics and Motor Control Priors

- Priors about gravity : objects fall down
- Mask: rotate by 90 degrees

Human Prior: Physics and Motor Control Priors

- Priors about gravity : objects fall down
- Mask: rotate by 90 degrees
- Priors about muscle memory: left key makes player go left
- reversed the arrow key controls

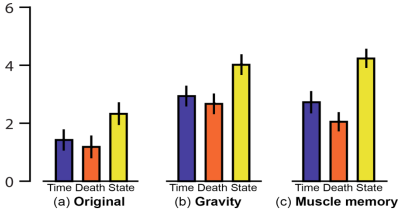
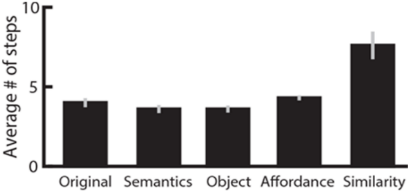


Figure: Performance of RL agent on various game manipulations. The RL agent performs similarly on all games except for the one without visual similarity

Comparing RL Agents: Controlling for change in complexity

- It is possible that humans take longer to solve the game because it is visually more complex
- But if RL Agents perform similarly with and without prior that means the modified games are not harder because of visual complexity.
- RL Agents show similar performance
- except : visual similarity
- "due to to the use of convolutional neural networks that implicitly impose the prior of visual similarity rather than simply due to the change in visual complexity."



(e) Performance of RL agent

Wrong Prior

- rewards in hidden locations
- human participants immediately assume that princess is the goal and do not explore the free space containing hidden rewards.
- a random agent ends up obtaining almost four times more reward than human players
- under-constrained exploration

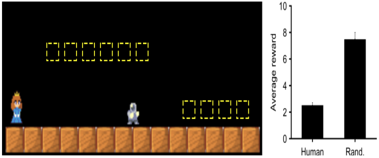


Figure: Prior information constrains human exploration.

Possible Solutions

- Goal: Algorithms that require fewer interactions with the environment: sample efficient
- incorporate prior initially instead of learning from scratch.
- learn priors through continual learning