INVESTIGATING HUMAN PRIORS FOR PLAYING VIDEO GAMES

Anonymous authors
Paper under double-blind review

ABSTRACT

Deep reinforcement learning algorithms have recently achieved impressive performance in playing video games. However, they require orders of magnitude more time than average human players to achieve the same performance. What makes humans so good at solving and figuring out these seemingly complex games? Here, we study one aspect critical to human decision making and problem solving – their use of strong priors (either learnt or inbuilt), that helps them to generalize and solve tasks faster (as opposed to learning from scratch). Through systematic investigation of human performance in video games, we develop a taxonomy of different forms of prior knowledge employed by humans that enables them to quickly solve video games. While common wisdom might suggest that prior knowledge about game semantics such as ladders are to be climbed, jumping on spikes is dangerous or the agent must fetch the key before reaching the door are crucial to human performance, we find that instead more general and high-level priors such as the world is composed of objects, object like entities are used as subgoals for exploration, and things that look the same, act the same are more critical. We hope that our findings will inspire the reinforcement learning community to make use of more structured representations for building more efficient and possibly human like agents.

1 INTRODUCTION

Consider the game shown in Figure 1.a. An agent playing this game receives a reward of +1 when the robotic sprite reaches the princess on top left. The agent receives no other reward and the game is reset if the agent steps on the fire or touches the angry purple sprite. We asked forty people to play this game (without communicating the reward structure of the game) and found that participants completed this game very quickly, with average time being just under a minute (refer to Figure 1.c). We then trained a state-of-the-art reinforcement learning (RL) agent (ICM-A3C; (Pathak et al., 2017)) on the same game and found that the RL agent took a significantly longer time to solve this game (close to four million steps).

What makes humans so efficient in solving this and other such games? Before we answer this question, now consider the game shown in Figure 1.b – which is just a simple modification of the previous game re-rendered with new textures. We again recruited forty people to play this game and found that human performance in this second game is much slower, with the average completion time being nearly two times that of the first game (Figure 1.c). In contrast, when we retrained the RL agent on this game, we found that the RL agent again took a long time close to 4 millions interactions to solve the game, and it was virtually unaffected in its performance.

What causes human performance to change so greatly between these two games? In the first game, without being given any information about the task on hand, the reader could have probably figured out that tiny robot sprite is the player and its goal is to reach the princess. The angry purple sprite and the fire objects are dangerous, the brick-like objects are platforms that the player can step on and the ladders can be used to reach the higher platforms. However, could the reader make such inferences in case of the second game? In the second game, it is hard for human players to decipher which object is the player sprite let alone what objects are the enemies, ladders, platform etc. This raises an interesting question, why does the difference in the two games not affect the performance of the RL agent? In contrast to human players, the RL agent comes with no prior knowledge about

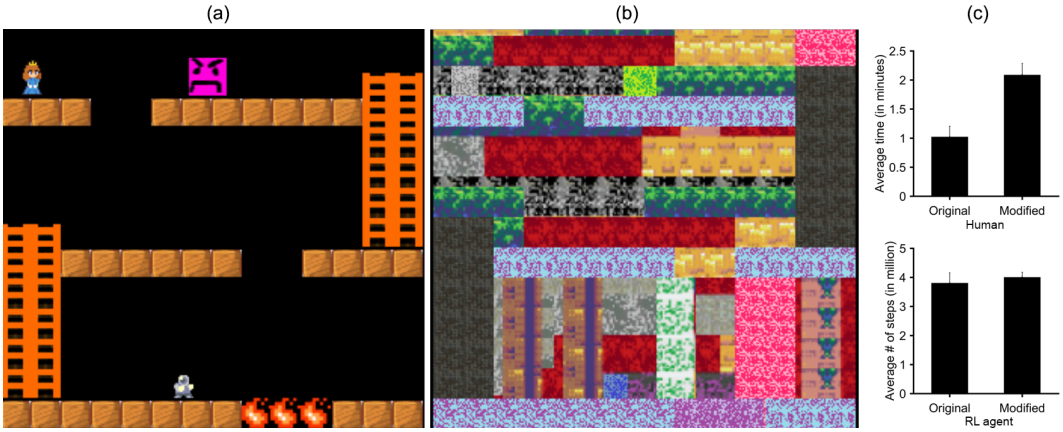

Figure 1: **Prior knowledge affects humans but not RL agents**. (a) A simple platformer game, (b) The same game modified by re-rendering the textures, and (c) human players and RL agent performance in the two games.

the physical world and both these games are more or less the same from the perspective of the agent. Thus, the agent needs to learn everything from scratch in both games and it takes a longer time to solve both these games.

This simple experiment highlights a link between the prior knowledge that people draw upon and the stark contrast that exists in the training time and efficiency between human players and RL algorithms ((Lake et al., 2016), (Tsividis et al., 2017)). Developmental psychologists have begun to document the prior knowledge that children draw upon in learning about the world ((Spelke & Kinzler, 2007) and (Carey, 2009)), however, two critical bottlenecks exist in this context. First, the field still lacks a complete understanding about various forms of prior knowledge that guide human behavior. Secondly, it is not yet clear as to how RL agents can either incorporate or learn about such prior knowledge. While some recent works have investigated this in the form of sub-goals and language grounding (Kulkarni et al. (2016), Narasimhan et al. (2017)), progress will be constrained till the field develops a better understanding of the kinds of structured prior knowledge humans employ.

In this paper, we primarily investigate and ultimately define a taxonomy of the different kinds of prior knowledge humans employ in order to successfully play video games. We chose video games as the platform of our investigation because it is easy to systematically change the game to include or mask different kinds of knowledge, run large-scale human studies, and video games such as ATARI are a popular choice in the reinforcement learning community. A major contribution of our paper is the provision of a benchmark and a taxonomy of various kinds of object priors people employ when solving video games. While common wisdom suggests that prior knowledge of the form that ladders are to be climbed, keys are used to open doors, jumping on spikes is dangerous is what enables humans to quickly solve the game, our work shows that such knowledge is not as critical as more general and high-level prior knowledge of the form that the world is composed of objects and things that look the same behave the same. Interestingly, these general object priors are learned by human babies when they are as young as 2 months old, implying that prior knowledge that is central to our core cognition plays a critical role in successful human gameplay. Our hope is that such a systematic investigation about the kinds of prior knowledge that make humans good at reinforcement learning tasks will inspire the reinforcement learning community to make use of more structured (and human like) representations in their systems.

2 METHOD

To investigate the aspects of visual information that enable humans to efficiently solve video games, we designed an elaborate browser based platform game consisting of a human sprite that can be controlled, platforms, ladders, slimy pink sprites that kill the agent, spikes that are dangerous to

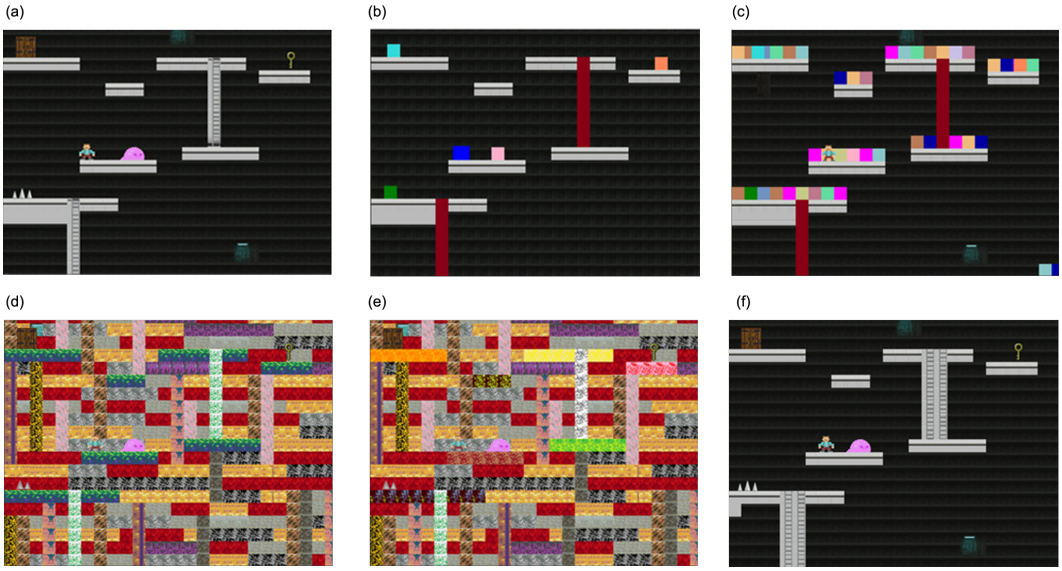

Figure 2: **Various game manipulations**. (a) Original version of the game. (b) Game with masked objects to lesion semantics prior . (c) Game with masked objects and distractor objects to lesion concept of object. (d) Game with background textures to lesion affordance prior. (e) Game with background textures and different colors for all platforms to lesion similarity prior. (f) Game with modified ladder to hinder people's prior about ladder properties.

jump on, a key and a door (see Figure 2). The human sprite can be moved with help of arrow keys and the agent obtains a reward of +1 when it reaches the door after taking the key and the game terminates. The game is reset whenever the agent touches the enemy, jumps on the spike or falls below the lowermost platform. We made this game to resemble the exploration challenges faced in the classic ATARI game of Montezuma's revenge that has proven to very challenging for state of deep reinforcement learning techniques (Bellemare et al., 2016; Mnih et al., 2015).

We systematically created different versions of this game by re-rendering various entities such as ladders, enemies, keys, platforms etc. using alternate textures (see Figure 2). These textures were chosen to mask various forms of prior knowledge that are described in the experiments section. Our experiment style draws inspiration from the neuroscience literature wherein researchers study aspects about the human brain by performing lesion studies (Müller & Knight (2006), Shi & Davis (1999)). For the purposes of our experiment, since it was not possible to go directly inside people's brain in order to study the importance of various priors, we did the next best thing possible - mask those priors.

For each version of the game created, we quantified human performance by recruiting thirty participants from Amazon Mechanical Turk. Each participant was instructed to use the arrow keys to move and finish the game as soon as possible. No information about the goals or the reward structure of the game was communicated to the human subjects. Each subject was paid $1 for successfully finishing the game. The maximum time for allowed for playing the game was set to 30 minutes. For each human subject we recorded the (x, y) position of the player at every step of the game, the total time taken by the subject to finish the game (i.e. achieve the reward of +1) and the total number of deaths prior to finishing the game. We used this data to quantify people's performance.

3 Quantifying the importance of object priors

The first version of the game shown is shown in Figure 2(a) (Game link - `https://dry-anchorage-61733.herokuapp.com/`). From a single glance at the game, people can employ their prior knowledge to interpret that the game agent can climb on ladders, it is supported by platforms, the pink slimy sprite is dangerous, spikes are to be avoided and probably the goal of the game is to take the key to open the door. As expected such interpretations enable people to

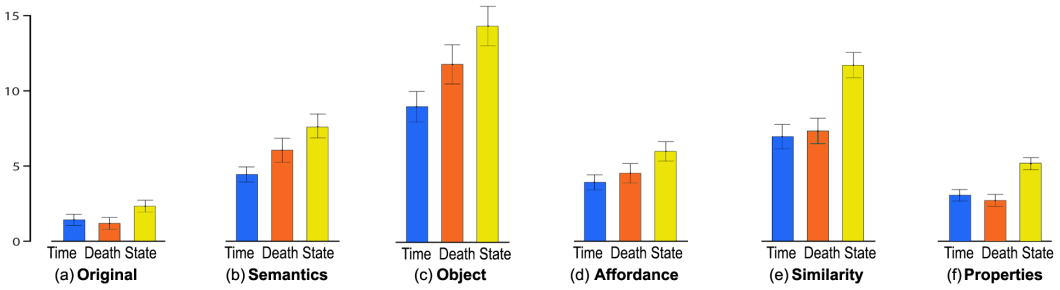

Figure 3: **Quantifying the influence of various object priors.** Blue bar shows average time taken by people to solve the various games, orange bar shows average number of deaths in the games, and yellow bar shows number of unique states visited by players in the various games. For visualization purposes the number of deaths is divided by 2 and the number of states is divided by 1000 respectively.

quickly solve the game. Figure 2(a) shows that the average time taken to complete the game is 1.4 minutes and the average number of deaths (shown by orange bar) and game states visited by the agent (shown by yellow bar) are quite small.

3.1 SEMANTICS

Intuition suggests that removing the semantic information from the game should lead to worse human performance. In order to study how important semantics are, we re-rendered the objects and ladders with blocks of uniform color as shown in Figure 2(b) (Game Link: `https://boiling-retreat-38802.herokuapp.com/`). In the current form all objects look alike and there is no apriori information about object types that could be used to guide exploration. As shown in Figure 3(b), people take 4.4 minutes on average to complete this game, which is approximately three times the time taken in full version. Furthermore the average number of deaths is 12 and humans explore significantly larger number of states (p-value < 0.05) as compared to the game shown in Figure 2(a).

While the results until now demonstrate that absence of semantics hurts human performance, it is unclear how humans make use of semantics. One hypothesis is that a critical role of semantics is to allow people to infer the latent reward structure of the game. If this indeed is the case then in the full version of the game, people should first visit the key and then go to the door, while in the version without semantics people should not exhibit any such bias. To investigate this, we computed the number of participants that reached the key before the door in both game conditions. We found that in the original game where key and door both were visible all thirty participants reached the key first, while in the version where semantics were ablated only ten out of thirty participants reached the key before the door (see Figure 4(a)). In order to perform a more detailed analysis we further investigated the time taken by participants to go from the key to the door. Because the average time taken by participants is different in both conditions we normalized the time taken by each participant to go from the key to the door by the total amount of the time that participant spent in the game. Results in Figure 4(b) show that people spend significantly more amount of time to reach the door when the semantics are hidden. This finding further provides evidence that in absence of semantics, people are unable to infer the reward structure and consequently significantly increase their exploration (see supplementary materials for a more detailed analysis).

In the setup described above we hid semantic information by recoloring objects with plain colors. An alternate mechanism to manipulate the semantic prior is by reversing the semantic of different entities (i.e. objects people associate as good were bad, and vice versa)) We created this version by replacing the pink enemy and spikes by coins and ice-cream sprite respectively which have a positive connotation, the ladder by fire, the key and the door by fire and spikes which have negative connotations. We found that performance of particiants in this condition was similar to the condition shown in Figure 3(b). Detailed analysis are reported in the supplementary materials here - `https://sites.google.com/view/human-rl/home`.

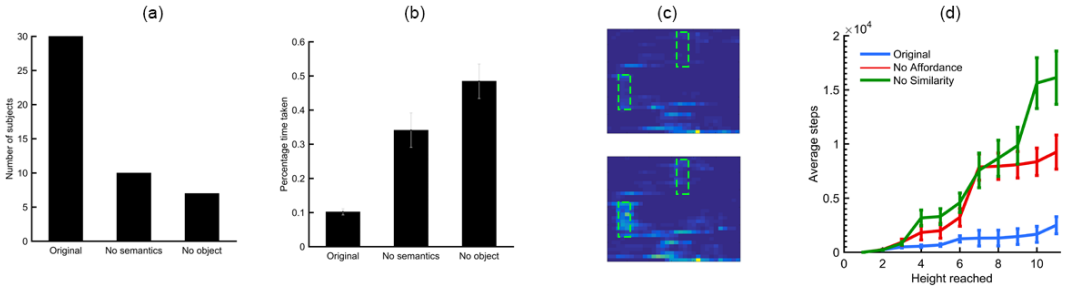

Figure 4: **Change in behavior upon lesion of various priors**. (a) Graph comparing number of participants that reached the key before the door in the original version, game without semantics, and game without object prior. (b) Graph showing amount of time taken by participants to reach the door once they obtained the key. (c) Heatmap comparing exploration trajectories of participants in full version of game (top) with respect to game with zigzag ladders (bottom). Ladders are highlighted via the green dashed boxes. (d) Graph showing average number of steps taken by people to reach various vertical levels in original version, game without affordance, and game without similarity.

3.2 Objects as Subgoals for Exploration

If the object semantics are as important for exploration as shown in the previous section, it would suggest that we need to equip deep reinforcement learning agents with semantic knowledge of a large number of objects that we encounter in every day life to make their exploration efficient. Is this really the case? Various studies in infant development suggest that even though babies might not recognize the semantic category of the object they still play with objects (Gopnik et al., 1999; Smith & Gasser, 2005; Spelke & Kinzler, 2007). This points to a speculation that people might have a more general prior knowledge about existence of objects and that objects are interesting entities to be explored. In the version of the game shown in Figure 2(b) the objects are distinctively visible and attract attention and it is possible that human subjects are instinctively biased to move towards them. In order to test this hypothesis, we modified the game to cover each space on the platform with a block of different color to hide where the objects are (see Figure 2(c), Game link - `https://high-level-1.herokuapp.com/`). Note that most colored blocks are placebos and do not correspond to any object and the actual objects have the same color and form as in the previous version of the game without semantics (i.e. Figure 2(b)). If the prior knowledge about entities that look like objects are interesting to explore is critical, this manipulation in the game structure should lead to a significant change in human performance.

Results shown in Figure 3(c)show that masking where objects are leads to drastic deterioration in performance and the average time taken by human subjects to solve this game rises sharply to 9 minutes and number of deaths to nearly 24 (Figure 3(c)). In absence of knowledge of objects, participants explore six times as many game states which is much larger than the number of states explored when only semantic information was removed. These results suggest that knowing that world consists of objects and objects can be used as subgoals for exploration is a very important prior used by humans. In our setup it is possible to further investigate this hypothesis by measuring the time between the agent goes to the object which is supposed to be the key and object which is supposed to be the door. When the location of objects is not known to the player, this time should be larger as compared to the time taken by players when object locations are known but there semantics might be unknown. The results shown in Figure 4(b) confirms this.

3.3 Affordance

In the previous versions of the game while we manipulated with priors about existence and type of objects, it was obvious for a human to infer what parts were platforms that could support the agent sprite and what parts of the game constituted free space. While the ladders did not have the expected texture the connectivity pattern of the platform hinted at that red regions in games shown in Figure 2(b,c) could be climbed to reach other platforms. This knowledge about what do with certain entities (i.e. ladders are to be climbed, platforms are there to support the agent) is referred

to as the affordance of the entity. In our next study we systematically removed the affordance prior while maintaining objects and their semantics. Note that we have purposefully constructed a difference between entities such as key, door, enemy, spike which cannot directly be used by the agent but convey the task structure and entities such as platforms, ladders and free space which do not necessarily convey the reward structure but are used to explore the environment.

We manipulated the affordance prior by covering the background with random textures and using textures on ladders and platforms to be visually distinctive (Refer to Figure 2(d), Game link: `https://fierce-sierra-47669.herokuapp.com/`). Results in Figure 3(d) show that manipulation of affordance information increases exploration and time taken by humans to solve the game as compared to the original game, but affordance appears to be less critical than object semantics in our setup.

3.4 THINGS THAT LOOK SIMILAR BEHAVE SIMILAR

In the previous game, although we manipulated information about affordances, once the player realizes that he can stand on a certain texture and climb up a certain texture, it is easy to use similarity in terms of color/texture to identify the platforms and ladders. Similarly in the version of the game without semantics (Figure 2(b)), visual similarity can be used to identify alternate enemies and spikes. This suggests that a general prior knowledge of the form that things that looks the same, act the same might guide fast exploration even in environments where semantics or affordances might be hidden.

We tested this hypothesis by modifying the no-affordance game such that the none of the platforms and ladders had the same visual signature. (Figure 2(e), Game link: `https://high-level-3.herokuapp.com/`). This prevented the human subjects from using the similarity prior and we found out that this makes them significantly inefficient at game play. The average time taken by players to solve the game was 7 minutes with number of deaths increasing to nearly 15, and the exploration was greatly increased as well (Figure 3(e)). The time taken to solve this game is significantly more ($p < 0.05$) than the original game, the one with no semantics (Figure 2(b)) and the game without affordances (Figure 2(d)) and not statistically different from the game where objects were hidden (Figure 2(c);$p = 0.1$). The number of deaths in this was significantly different than all versions of the game considered until now. This suggests that notion of similarity is the second most important object prior in human gameplay after the knowledge of directing exploration towards objects.

In order to gain insight into how this prior knowledge effects humans, we investigated the exploration pattern of human players. In the game when all information is visible we expect that progress of humans would be uniform in time. In the case when affordances are removed, the human players would initially take sometime to figure out what visual pattern corresponds to what entity and then quickly make progress in the game. Finally, in the case when the similarity prior is removed, we would expect human players to be unable to generalize any knowledge across the game and take large amounts of time exploring the environment even towards the end. We investigated if this indeed was true by computing the time taken by each player to reach different vertical distances in the game for the first time. Note that the door is on the top of the game, so the moving up corresponds to getting closer to solving the game. The results of this analysis are shown in Figure 4(d). The x-axis shows the height reached by the player and the y-axis show the average time taken by the players. As the figure shows, the results confirm our hypothesis.

3.5 HOW TO INTERACT WITH OBJECTS

Until now we have analyzed prior knowledge used by humans to interpret the visual structure in the game. However, interpretation of visual structure is only useful if the human player understands what to do with the interpretation. For example, once a player recognizes an object (i.e. door, monster, ladder), they seem to possess prior knowledge about how to interact with that object - monsters can be avoided by jumping over them, ladders can be climbed by pressing the up key repeatedly etc. Deep reinforcement learning agents on the other hand do not possess such priors and must learn how to interact with objects by mere hit and trial.

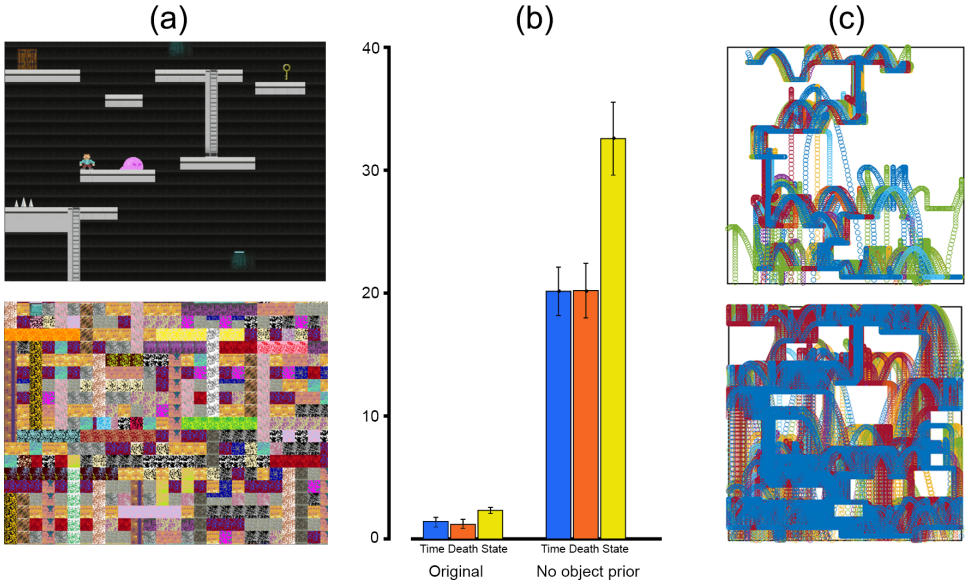

Figure 5: **Masking all object priors drastically affects human performance**. (a) Full version of the game (top) and version of game without any object priors (bottom). (b) Graph depicting difference in people's performance for both the games. (c) Exploration trajectory for full version (top) vs no object prior version (bottom).

To test how critical is such prior knowledge, we created a version of the game in which the ladders couldn't be climbed by simply pressing the up key. Instead, the ladders were zigzag in nature and in order to climb the ladder players had to press the up key, followed by alternating presses between the right and left key. Note that the ladders in this version looked like normal ladders, so players couldn't infer the properties of the ladder by simply looking at them (see Figure 2(f), game link: `https://calm-ocean-56541.herokuapp.com/`). As shown in Figure 3(f), changing the property of the ladder increases the time taken by players to solve the game. The time spent by the agent in different parts of the game visualized in this original game (top row of Figure 4(c)) and this game reveals (bottom row of Figure 4(d)) that humans spend significantly more amount of time in the first ladder in the modified version of the game. However, once they learn about how to use the ladder they are able to quickly climb the second ladder.

Finally, even though the average number of deaths increased in this version, we note that the number is still lower than number of deaths in game versions without semantics and without affordance ($p < 0.05$). Critically, the time taken by players to solve this version of the game is significantly different than time taken to solve the version without semantics with $p < 0.05$ again. This finding suggests that while prior knowledge about object properties plays a critical role in human gameplay, knowledge about semantics and affordances may be more important than this prior.

4 Creating a taxonomy of object priors

In the previous section, we quantified the importance of various object priors in the context of human gameplay. What happens to human gameplay when they don't have access to all of those object priors? To answer this question, we modified the game by re-rendering the objects, platforms, and ladders to hide all information about objects, semantics, affordance, and similarity (refer to Figure 5.a). This game is in a way close to how the RL agent perceives a normal game screen - without having any notion of objects, semantics, affordances, and similarity. To quanitify human performance, we again recruited participants from Mechanical Turk to play this game ($n = 26$, pay = \$2.25 to encourage completion of the game).

As shown in Figure 5.b, the performance of humans becomes extremely poor in this version of the game. The average time taken to solve the game increases to 20 minutes and the average number

of deaths rises sharply to 40. Remarkably, the exploration trajectory of humans is now almost completely random (refer to Figure 5.3.c) with the number of unique states visited by the human players increasing by a factor of 9. In this version, human players are now performing actions and visiting states that they wouldn't do otherwise. We also experienced an extremely high drop out rate for this version of the game with many participants noting that the only reason they were able to solve the game was due to their memory.

Thus, despite the fact that humans come equipped with powerful optimization tools and algorithms, weakening their representation by removing prior information makes problem solving extremely hard for humans. If we consider that this version only removes object priors while preserving priors related to physics and motor control, then these results become even more significant. If we were to create a game that reverses physics and motor, we do not foresee human players even attempting to solve the game. As a quick test, we designed one such game but had an extremely hard time trying to solve the game (despite being the creators of the game and playing previous version of the games 100s of times). How then can we expect RL agents to solve such complex games at the efficiency of humans? While there is no doubt that deep RL agents will improve with better algorithms and better computational resources (which the field has made remarkable progress on), these results suggest the value of incorporating prior knowledge towards achieving these goals.

How then should RL agents go about incorporating these prior knowledge? We next relate our findings to literature from developmental psychology to suggest a possible taxonomy and ranking of these object priors in the hope of answering this question.

Human infants as young as 2 months old develop primitive notions about objects such as coherence and consistency (Aguiar & Baillargeon (1999), Spelke (1990)). Thus, concepts of objects are at the core of our cognition and helps us make reason efficiently in various situations. Here, in our game, players come with a very general and high level prior about objects i.e. there are objects in the world and this prior in itself is the most important prior in guiding human gameplay (Section 3.2).

By the time babies are 3-5 months old, they start exhibiting categorization behavior based on similarity and familiarity (Mandler (1998), Mareschal & Quinn (2001)). Thus, the notion of similarity itself forms a central part of human cognition. In line with this, our findings show that prior information regarding similarity is removed, human gameplay is affected significantly (Section 3.4).

While very young infants are able to perceive simple shapes, the ability to recognize individual objects rapidly and accurately emerges comparatively late in development (usually by the time babies are 18-24 months old Pereira & Smith (2009)). Similarly, while young infants exhibit some knowledge about affordances early during development, the ability to distinguish a walkable step from a cliff emerges only by the time they are 18 months old(Kretch & Adolph (2013)). Consistent with these results, our findings show that while prior information about semantics and affordances play an important role in human gameplay, they are both less important than general priors such as similarity and concept of object (Section 3.1 and Section 3.3).

Taken together, this suggests that the earlier a prior is learned during the course of development, the more critical that prior is during gameplay. Based on these results, we suggest a possible taxonomy and ranking of object priors in Figure 6 . We here put object properties at the bottom as in the context of our problem, object properties about specific objects can be only learnt once recognition is performed.

The significance of this figure to the RL community is pretty straightforward – even when humans encounter a new, complex environment in which they have no prior information about semantics, affordance, or object properties, they are still able to reason efficiently in those environments. This is due to the fact that they still have notion of general priors that they learned very early on during development and these general priors help them perform guided exploration. We believe that incorporating such general priors to RL agents can not only lead to faster training time, but it will help RL algorithms generalize better in newer and unknown environments.

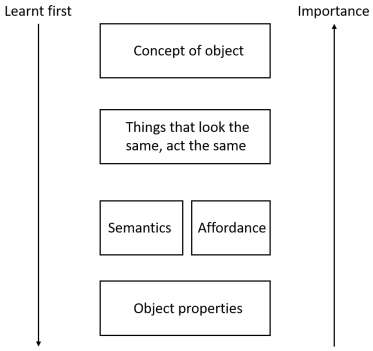

Figure 6: **Taxonomy of object priors.** The earlier an object prior is obtained during childhood, the more critical that object prior is in human problem solving in video games.

5 WHEN PRIOR KNOWLEDGE CAN SOMETIMES LEAD TO SUBOPTIMAL REWARDS.

For many interesting real world tasks, for pragmatic reasons it is often only possible to provide agents with a terminal reward when they succeed and they receive no external rewards otherwise. Success in such scenarios critically depends on the agent's ability to explore its environment and then quickly learn from its success (i.e. exploitation). While understanding what enables an agent to efficiently exploit is an interesting question, without a good exploration strategy no exploitation is possible. It therefore naturally follows that agents that can efficiently explore their environment will be good at completing tasks with sparse rewards. In this vein, our results demonstrate the importance of prior knowledge in helping people explore efficiently in these sparse reward environments.

However, that being said, being equipped with strong prior knowledge may not be beneficial with regards to reward optimization in all kinds of environments. To illustrate this, we created a short game that simply consisted of a player and a princess at a short distance away from the player (7.a). Unknown to the participants, the game consisted of 10 hidden rewards (shown in yellow for illustration purposes) and the participants were given a bonus for discovering them. Upon entering the game, the players see the princess and infer that as the goal and immediately reach her, thereby terminating the game. As shown in (7.b), human players do not explore this environment and end up with suboptimal rewards. In contrast, a random agent (30 seeds of episode count =1 to simulate human experiments) ends up obtaining almost 4 times the rewards than human players.

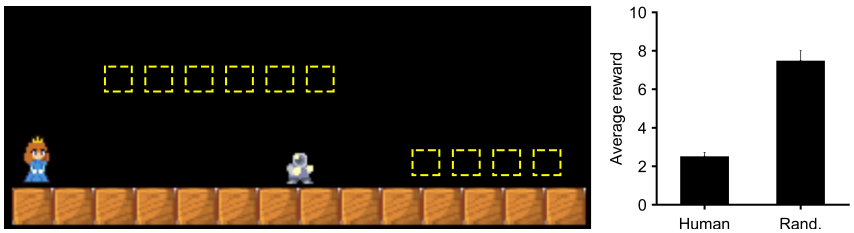

Figure 7: **Prior information constrains human exploration.** (Left) A very simple game with hidden rewards (shown in dashed yellow). (Right) Average rewards accumulated by human players vs a random agent.

Research in developmental psychology has also demonstrated such instances wherein children have been shown to be better learners than adults in some cases (Lucas et al. (2014)). Thus, while incorporating prior knowledge in RL agents has many potential benefits, it is also important to consider if that could lead to inflexibility in an algorithm leading to inefficient exploration.

6 CONCLUSION

While there is no doubt that the recent performance achieved by current deep RL algorithms on complex environments is beyond impressive, there is a lot to learn from human cognition if we want RL agents to reduce their training time as well as perform efficiently in sparse reward settings. One of the most compelling arguments for this is the ease at which humans are able to learn from limited data and generalize their knowledge to new domains. A key reason behind this is that whenever people encounter a new problem, they are informed by their rich experiences and prior knowledge about the world. Our work takes one of the first steps to signify the importance of prior knowledge and understanding how prior knowledge makes humans good at reinforcement learning. We believe that these results and data set can serve as a benchmark for researchers thereby laying down the foundations for future RL algorithms that intend to incorporate such prior knowledge in agents.

REFERENCES

Andréa Aguiar and Renée Baillargeon. 2.5-month-old infants' reasoning about when objects should and should not be occluded. *Cognitive psychology*, 39(2):116–157, 1999.

Marc Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Remi Munos. Unifying count-based exploration and intrinsic motivation. In *NIPS*, 2016.

Susan Carey. *The origin of concepts*. Oxford University Press, 2009.

Alison Gopnik, Andrew N Meltzoff, and Patricia K Kuhl. *The scientist in the crib: Minds, brains, and how children learn.* 1999.

Kari S Kretch and Karen E Adolph. Cliff or step? posture-specific learning at the edge of a drop-off. *Child Development*, 84(1):226–240, 2013.

Tejas D Kulkarni, Karthik Narasimhan, Ardavan Saeedi, and Josh Tenenbaum. Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation. In *Advances in Neural Information Processing Systems*, pp. 3675–3683, 2016.

Brenden M Lake, Tomer D Ullman, Joshua B Tenenbaum, and Samuel J Gershman. Building machines that learn and think like people. *Behavioral and Brain Sciences*, pp. 1–101, 2016.

Christopher G Lucas, Sophie Bridgers, Thomas L Griffiths, and Alison Gopnik. When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, 131(2):284–299, 2014.

Jean M Mandler. Representation. 1998.

Denis Mareschal and Paul C Quinn. Categorization in infancy. *Trends in cognitive sciences*, 5(10): 443–450, 2001.

Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 2015.

NG Müller and RT Knight. The functional neuroanatomy of working memory: contributions of human brain lesion studies. *Neuroscience*, 139(1):51–58, 2006.

Karthik Narasimhan, Regina Barzilay, and Tommi Jaakkola. Deep transfer in reinforcement learning by language grounding. *arXiv preprint arXiv:1708.00133*, 2017.

Deepak Pathak, Pulkit Agrawal, Alexei A Efros, and Trevor Darrell. Curiosity-driven exploration by self-supervised prediction. *arXiv preprint arXiv:1705.05363*, 2017.

Alfredo F Pereira and Linda B Smith. Developmental changes in visual object recognition between 18 and 24 months of age. *Developmental science*, 12(1):67–80, 2009.

Changjun Shi and Michael Davis. Pain pathways involved in fear conditioning measured with fear-potentiated startle: lesion studies. *Journal of Neuroscience*, 19(1):420–430, 1999.

Linda Smith and Michael Gasser. The development of embodied cognition: Six lessons from babies. *Artificial life*, 2005.

Elizabeth S Spelke. Principles of object perception. *Cognitive science*, 14(1):29–56, 1990.

Elizabeth S Spelke and Katherine D Kinzler. Core knowledge. *Developmental science*, 10(1):89–96, 2007.

Pedro A Tsividis, Thomas Pouncy, Jacqueline L Xu, Joshua B Tenenbaum, and Samuel J Gershman. Human learning in atari.". In *The AAAI 2017 Spring Symposium on Science of Intelligence: Computational Principles of Natural and Artificial Intelligence*, 2017.