Arachnophobia Exposure Therapy using Experience-Driven Procedural Content Generation via Reinforcement Learning (EDPCGRL)

Athar Mahmoudi-Nejad, Matthew Guzdial, Pierre Boulanger

Department of Computing Science, University of Alberta
athar1@ualberta.ca, guzdial@ualberta.ca, pierreb@ualberta.ca

Abstract

Personalized therapy, in which a therapeutic practice is adapted to an individual patient, leads to better health outcomes. Typically, this is accomplished by relying on a therapist's training and intuition along with feedback from a patient. While there exist approaches to automatically adapt therapeutic content to a patient, they rely on hand-authored, pre-defined rules, which may not generalize to all individuals. In this paper, we propose an approach to automatically adapt therapeutic content to patients based on physiological measures. We implement our approach in the context of arachnophobia exposure therapy, and rely on experience-driven procedural content generation via reinforcement learning (EDPCGRL) to generate virtual spiders to match an individual patient. In this initial implementation, and due to the ongoing pandemic, we make use of virtual or artificial humans implemented based on prior arachnophobia psychology research. Our EDPCGRL method is able to more quickly adapt to these virtual humans with high accuracy in comparison to existing, search-based EDPCG approaches.

Introduction

Experience-driven Procedural Content Generation (EDPCG) is a PCG framework that modifies content to optimize a user's experience. Although EDPCG was developed for games, it can be applied to other HCI domains that require automated customized content (e.g. recommender systems) (Yannakakis and Togelius 2011). We argue EDPCG can also be a useful framework for computer assisted-therapy.

We can divide computer-assisted therapy into two groups: non-adaptive and adaptive. The non-adaptive approach provides predesigned content for all users, which is not ideal since individuals benefit from individualized treatment (Zahabi and Razak 2020). Adaptive approaches generally structure this as a selection problem: choosing between pre-existing content to better match an individual's treatment needs, resulting in better health outcomes (Zahabi and Razak 2020). Current adaptive computer-assisted therapy is mainly therapist-guided or rule-based. The former requires therapist intervention which is time-consuming, burdensome, and requires specialized training. The latter modifies the therapeutic content based on predefined rules, which are brittle and

cannot account for all possible individuals (Abdessalem and Frasson 2017; Heloir et al. 2014). In contrast, an EDPCG framework could be used in computer-assisted therapy as an automatic content generation tool that adapts to satisfying individuals' therapeutic needs.

There are only a handful of studies that have used EDPCG for physical rehabilitation, e.g., motor rehabilitation (Dimovska et al. 2010), amblyopia (Correa et al. 2014), and upper limb rehabilitation (Hocine et al. 2015). In these studies, the Player Experience Model (PEM) in the EDPCG framework is gameplay-based (player performance). Gameplay-based PEM assumes that the player's internal state can be derived from the way a player plays a game. However, according to Yannakakis and Togelius (Yannakakis and Togelius 2011), gameplay-based PEM is a low-resolution model of the player's experience, which is not ideal for psychological rehabilitation. Further, less attention has been paid to using EDPCG for psychological rehabilitation, e.g., mental health (Badia et al. 2018). We draw on physiological measures in our PEM, which more closely correspond to a player's internal state, especially for measuring stress.

This paper presents an EDPCG framework for arachnophobia treatment leveraging a Experience-driven Reinforcement Learning-based PCG (EDPCGRL) content generator. The RL agent dynamically generates spiders in order to induce a desired stress level. The framework's goal is to keep players within a defined physiological state to allow for more effective exposure therapy. Exposure therapy is a therapy technique for treating anxiety disorders in which an individual is gradually exposed to the anxiety source. This paper introduces a new research area: EDPCGRL. We investigate the application of an EDPCGRL system to computer-assisted therapy. The contributions of our framework are: 1) Demonstrating the feasibility of a cognitive-based EDPCGRL approach for rehabilitation; 2) Demonstrating the first instance, to the best of our knowledge, of PCGRL outperforming search-based PCG.

Related Work

VR for Arachnophobia

Virtual reality (VR) immerses individuals in graphical computer-generated environments. VR has been found to be effective in exposure therapy for specific phobias, in-

Study	Reward Function	Policy	Focus	Case Study
(Guzdial, Liao, and Riedl 2018)	Usability and helpfulness of the RL's suggestion	Deep Q-Learning	Partner in co-creative level design	Super Mario Bros.
(Nam and Ikeda 2019)	Difficulty-based evaluation function	Deep Q-Learning, Deep Deterministic Policy Gradient	Generate diverse stages	Implemented platforms (battle and non-battle recovery sections)
(Lopez, Ashour, and Tucker 2019) (López et al. 2020)	Efficiency and functionality of the generated layout	Proximal Policy Optimization	Placement of virtual objects	Manufacturing facility
(Cunningham et al. 2020)	Efficiency and functionality of the generated layout	Proximal Policy Optimization	Placement of virtual objects	Manufacturing facility, Grocery store
(Khalifa et al. 2020)	Playable levels and satisfying some rules	Proximal Policy Optimization	2D maps and game element placement	Binary, Zelda, Sobokon
(Gisslén et al. 2021)	Based on the solver's performance	Adversarial RL	Environment (racing tracks, platforms, paths)	Racing game
(Susanto and Tjandrasa 2021)	Comparison between achieved and desired goal	Deep Q-Learning	Level generation	Bit-flipping, Lights out, Leaping Frog, Google chrome's Dinosaur game
(Earle et al. 2021)	Closeness of the generated level to target metrics	Proximal Policy Optimization	Level generation	Binary, Zelda, Sokobon, SimCity, micro-RCT

Table 1: Summary of prior PCGRL work.

cluding fear of heights (Freeman et al. 2018), fear of spiders (Arachnophobia) (Côté and Bouchard 2009) and other anxiety disorders (Maples-Keller et al. 2017). In exposure therapy, a subject is gradually exposed to a feared situation or object in a safe environment, leading to desensitization and a healthier response. While we are not dismissing non-adaptive VR for arachnophobia such as (Shiban et al. 2015; Miloff et al. 2019), we focus on adaptive approaches, because they better fit subjects' different needs. For example, Kritikos et al. (Kritikos, Alevizopoulos, and Koutsouris 2021) defined rules to change a spider's appearance and pattern of behaviour (e.g., size and velocity of the spider, probability of walking towards the user, etc.) to induce a desired level of anxiety in a subject. The level of anxiety is calculated based on normalized electrodermal activity (EDA [1]) changes. Instead of using predefined rules based on EDA changes, our framework adapts a spider using PCGRL.

PCG

Procedural Content Generation (PCG) automatically generates content using algorithms. Traditional PCG approaches require hand-authored knowledge, such as constructive (Shaker, Togelius, and Nelson 2016), search-based (Togelius et al. 2011), and constraint-based methods (Smith and Mateas 2011). To address this limitation, researchers started applying machine learning (ML) methods to PCG (Summerville et al. 2018); however, because they are primarily supervised learning methods, they require a pre-existing dataset. Our framework is based on PCGRL to automatically generate new content without the need of a dataset.

---

[1]The EDA measures the variations in the skin's electrical conductance due to increases in the activity of sweat glands

PCG for Rehabilitation

PCG has rarely been applied to rehabilitation. Dimovska et al. (Dimovska et al. 2010) used a constructive PCG generator in a ski-slalom game to place challenges according to a player's performance for motor rehabilitation. Correa et al. (Correa et al. 2014) developed an adaptive first-person shooter game using constructive PCG for amblyopia treatment based on the player's performance. Hocine et al. (Hocine et al. 2015) developed a game for upper limb rehabilitation through pointing tasks, i.e., reaching targets. They locate targets using Monte Carlo tree Search (MCTS) based on the user's performance, and generate the level with constructive PCG (i.e., choosing game entities). Badia et al. (Badia et al. 2018) developed a labyrinth game that adapts by estimating a user's emotions via physiological responses. The system promotes an emotional self awareness for more effective emotion regulation. A constructive PCG approach was applied to generate different graphical content. These works use a set of predefined content, assuming the subjects are known. Instead, our work assumes that the subjects are unknown; therefore, content needs to be generated and adapted dynamically.

PCGRL

PCG via Reinforcement learning (PCGRL) methods focus on applications where we do not have any pre-authored training data, but we do have an environment of possible content and a way of automatically evaluating that content. In contrast to supervised PCGML methods, no pre-existing data is required for PCGRL. Thus, it can be applied to situations like ours in which no data is available.

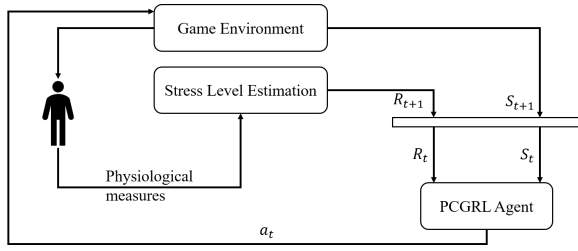We summarize prior PCGRL work in Table 1. The purpose of these projects is mainly entertainment or education.

Figure 1: Proposed system framework

.

	Attribute	Impact Factor	Possible Values
Movement Factors	(1) Locomotion	$\mu : 0.9$ $\delta : 0.15$	0: Standing 1: Human-like locomotion 2: Spider-like locomotion
	(2) Amount of movement	$\mu : 0.9$ $\delta : 0.15$	0: Slightly 1: Moderate 2: Too much
	(3) Closeness	$\mu : 0.4$ $\delta : 0.17$	0: Far away 1: In the middle 2: Very close
Appearance Factors	(4) Largeness	$\mu : 0.7$ $\delta : 0.16$	0: Small 1: Medium 2: Large
	(5) Hairiness	$\mu : 0.6$ $\delta : 0.21$	0: Without hair 1: With hair
	(6) Color	$\mu : 0.5$ $\delta : 0.20$	0: Gray 1: Brown 2: Black

Table 2: List of the Attributes and their Impact factors on fear (mean and standard deviation) based on (Lindner et al. 2019). We define ordinal possible values for each attribute.

The reward functions used in these works are based on different game-specific content. The application of a game-specific reward function, in our framework, would require assuming that the game-specific content reflects the players' cognitive state. However, the assumption is not necessarily true, and would be challenging to implement since it requires a sophisticated game design. Instead, based on Shaker (Shaker 2016), we define our reward function via human interaction with the system, which we argue is an essential factor in effectively using PCGRL for rehabilitation purposes.

Methodology

This section overviews our framework for adapting a game environment using PCGRL for exposure therapy based on user responses. We visualize our framework in Figure 1. We use arachnophobia as a case study for this framework. A subject interacts with a virtual spider, and their physiological responses are used to estimate their stress levels in real-time. A PCGRL agent modifies the virtual spider in order to reach a therapist-defined goal stress level. While we intend for the final version of this system to draw on prior work on estimating stress from physiological responses (Schmidt et al. 2018), we do not implement this part of the framework for this paper. In the following subsections, we describe each component of Figure 1 in detail.

Game Environment

For the environment, we envision a VR program with a generated 3D model of a spider. Our initial implementation does not require the completed VR environment and so we use a prototype version of the game environment, which is a simple representation of the spider given its specific attributes. We visualize three sets of attribute values for our prototype spider in Figure 2. The virtual spider is initially generated with random attributes, which are later adjusted by our ED-PCG approach. These attributes include movement-related and appearance characteristics of spiders based on Lindner et al.'s work (Lindner et al. 2019). In this study, spider-fearful individuals ($n = 194$) were asked to rate the characteristics of spiders based on their fear response. We overview this study in more detail in the next subsection.

Stress Level Estimation

Theoretically, players' physiological responses should allow for an estimation of their stress level (Schmidt et al. 2018). The estimated stress level is compared with a target stress level (explicitly identified for exposure therapy purposes by a therapist) to generate reward. Given the ongoing global health crisis, we calculate the stress levels of virtual subjects as functions of the spider's attributes in order to develop a simplified version of our framework as a proof of concept. Consequently, we generate various virtual subjects, i.e., each has its own stress level function corresponding to spider attributes, based on the Lindner et al. study (Lindner et al. 2019). The paper introduced seven spider attributes and measured their impact on their participants' overall amount of fear by asking the participants to quantitatively rate each spider attribute. We choose six attributes for spider generation as shown in Table 2. We drop the realness attribute since it indicates whether or not to use real spiders, and we focus on virtual spiders. For each attribute, we define 2-3 ordinal values.

The movement attributes: *locomotion*, *amount of movement* and *closeness*, denote how the spider moves (specifically the movement of the legs), how much it moves, and how close it gets to the subject respectively. The appearance attributes: *largeness*, *hairiness*, and *color*, denote the size of the spider, whether the spider is hairy or not, and the color of the spider, respectively.

We generate 100 virtual subjects as stress estimation functions, each with different responses to spider attributes. In (Lindner et al. 2019), an impact factor is associated with each spider attribute, denoting the effect of the attribute on the subjects' amount of fear. Based on the impact factors reported in the study, we model a normal distribution for each of the attributes as $N \sim (\mu_{a_i}, \delta_{a_i}), i \in \{1, ..., 6\}$; where a_i denotes the i-th attribute in Table 2, and μ_{a_i} and δ_{a_i} are derived from the i-th attribute's mean impact factor and subjects' fear variance of the attribute, respectively. We draw 100 random samples from each distribution to generate 100 virtual humans as our subjects. An example of stress for one virtual subject is: $1.37 \times (0.97 \times a_1 + 0.87 \times a_2 + 0.07 \times a_3 + 0.63 \times a_4 + 0.67 \times a_5 + 0.77 \times a_6)$; Where 1.37 is the coefficient to scale the stress level to the range 0-10 for

this virtual subject. This approach led to 100 unique virtual stress levels, which are still based on psychological study of real, spider-fearing humans. Thus, we expect there to be at least some psychological grounding to our results.

Markov Decision Problem

State   Each state represents a combination of the spider's attributes. We define for each attribute a range of possible values, which are listed in Table 2. The representation of state in time t can be represented as (Eq. 1):

$$S_t = \{a_{i,t}\}, i \in \{1, ..., 6\} \tag{1}$$

Where $a_{i,t}$ is the value of the i-th attribute in time t.

Action   The action is defined as increasing or decreasing one attribute at a time (Eq. 2).

$$A_t = \{a_{policy,t} \pm 1\} \tag{2}$$

Where $a_{policy,t}$ is the attribute chosen by the policy of ED-PCGRL for action in time t. The intuition behind this decision is that therapists gradually increase/decrease these attributes in exposure therapy to find reasonable values to produce their intended stress level.

Reward   The reward function is calculated based on a normal distribution $N(\mu, \delta)$, where μ is the target stress level, and $\delta = (MaxStress - MinStress)/2$, where $MaxStress$ and $MinStress$ are maximum and minimum values that the stress level can reach respectively. We scale the distribution to the range of $(-1, 1)$ such that the target stress level achieves a reward of 1, and the reward decreases as the stress level gets further from the target stress. The resulting reward function is shown in Eq. 3.

$$Reward = \frac{2e^{-0.5(\frac{x-\mu}{\delta})^2} - e^{-0.5(\frac{\alpha-\mu}{\delta})^2} - 1}{1 - e^{-0.5(\frac{\alpha-\mu}{\delta})^2}}$$
$$\alpha = \begin{cases} MaxStress, & \text{if } \mu < \delta. \\ MinStress, & \text{if } \mu \geq \delta. \end{cases} \tag{3}$$

where x is the current stress level.

Evaluation

We propose a framework for arachnophobia exposure therapy that can automatically adapt to subjects. For each subject, it generates spiders until it finds a spider (with specific features) that induces a subject's target stress level. This is a difficult task since it is suggested in (Vetter 2013) that people with arachnophobia respond to different aspects of spiders differently. We evaluate our proposed method in terms of how many times we present new spiders to a subject, denoted as *Spiders Presented*. This number should be as low as possible, because it may adversely affect the effectiveness of the treatment. This metric is equivalent to the number of times that an algorithm outputs a new spider.

At this stage, we evaluate our framework using virtual subjects (discussed in the methodology section) as a proof

of concept. This is because of the ongoing COVID-19 pandemic, which stopped us from conducting a human subject study. We generate 100 virtual subjects with a theoretical basis in real spider-fearing individuals. If our method outperforms our baseline methods using these virtual subjects, it will indicate that we can likely expect similar performance with real subjects.

Since there is no existing EDPCGRL framework to compare against, we compare our proposed method to EDPCG via search-based procedural content generation (SBPCG) methods. SBPCG is a traditional way to adapt content based on the assessment of users (Risi et al. 2015; Liapis, Yannakakis, and Togelius 2013). We briefly describe the baseline methods used in our evaluation and the implementation of our proposed method.

Genetic Algorithm: A Genetic Algorithm (GA) is our first baseline because it is a popular experience-driven SBPCG approach (Togelius et al. 2011). In our problem, the GA's chromosome is equal to the state defined in the previous section. The GA starts with a population that consists of a given initial state and its nine neighbour states (the population size is 10). We denote two states as neighbours if and only if they differ by only 1 value for a single attribute. Our approach selects the ten best chromosomes in terms of their fitness as its initial population. Based on Table 2, there are two chromosomes that could have eleven neighbours (locomotion=1, amount of movement=1, closeness=1, largeness=1, hairiness=0 or 1, color=1). The "average" initial state is one of these chromosomes. Therefore, we do not anticipate different results if we took all available neighbours. The GA then performs crossover, mutation, and selection until it reaches the termination condition. The crossover process uses a weighted sampling for picking two pairs of chromosomes based on fitness score. The pairs are swapped from the middle point (one child takes the first half of the attributes from one parent and the second half from the other, and the other child vice versa) to generate two new offspring. For mutation, a random attribute in the new offspring is changed with 0.1 probability to a random, valid. These operations do not hold the constraint of increasing/decreasing one attribute at a time, allowing for much larger steps through the search space. Finally, the selection process chooses ten chromosomes with the highest fitness values as the next population.

Greedy search: Our second baseline is a simple greedy search. This approach works best if the fitness function does not have local optima since it always exploits the best neighbour. It starts from the given initial state, searches its neighbours, and chooses the neighbour with the maximum fitness. Therefore, the fitness gradually increases (or stays steady) until the termination condition. This approach should perform best in terms of the minimum number of *Spiders Presented* if our problem is simple enough.

Random Search: Our final baseline is a random search. It also starts from the same initial state as previous baselines. In each step, it randomly selects an action (defined in the previous section) and goes to a new state. It keeps searching

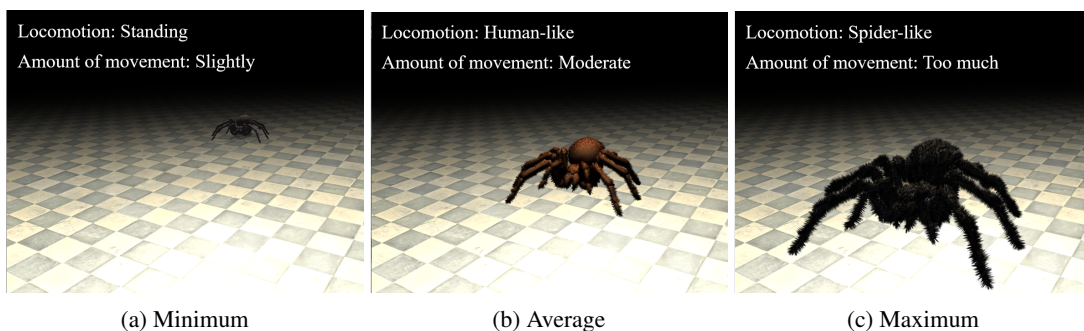|		
(a) Minimum	(b) Average	(c) Maximum

Figure 2: Generated spiders with three initial attributes from our initial prototype.

until it reaches the termination condition. We include random search to investigate the importance of exploration in this domain.

PCGRL (Our method): This is our method, which optimizes the spiders' attributes using RL. We employ the tabular Q-learning algorithm and epsilon-greedy ($\epsilon = 0.05$) for our action selection policy (Sutton and Barto 2018). We initialize a Q-table with either random or zero values that store the values of state-action pairs, and update the values in each iteration.

These methods all use the same fitness function: the reward function for the PCGRL agent. We use the same initial state and termination criteria for all the methods for a fair evaluation. We define these as:

- Initial State: We use three initial states for our evaluation: in each state, we set the spider's attributes to either the maximum, minimum, or the average values within each attribute's range (visualized in Figure 2).

- Termination: These algorithms terminate if they reach one of these criteria: achieve a state with maximum fitness or run for 100 iterations.

We run each algorithm 10 times for each virtual subject and for each of the three initial states. We apply these approaches for every goal stress level from 1 to 9. In total, each algorithm is executed 27000 times.

Result

We define three stress categories: low ($[1, 3]$), moderate ($[4, 6]$), and high ($[7, 9]$). We calculate the average value of *Spiders Presented* for our 100 virtual subjects for each initial state and each stress category. We repeat this process 10 times and report the averages in Table 3.

Spiders Presented in the table denotes the number of new spiders showed to a virtual subject on average when the approach is successful. The *Accuracy* shows the percentage of times the method could find a spider that provokes the target stress level out of the ten attempts on average. The *Spiders Presented* result is not taken into account if a run had *Accuracy* less than 75%. For example, for the *Max* initial state and *Low* stress goal, the *RL_Zero* has the least number of Spiders Presented, but since the Accuracy is 70.93% (less

than 75%), we did not consider it to be the best result. In each row, the method that outperforms other methods according to the *Spiders Presented* metric is shown with bold text. The table show that our proposed method outperformed the baseline methods for almost all of the target stress levels with various initial states. We note that we do not expect these results to necessarily generalize to all hyperparameter values. A hyperparameter sweep of the approaches is not in the scope and within the page limitation of this paper.

Two-tailed, paired-samples t-tests ($p < 0.05$) were performed to compare the mean *Spiders Presented* metric across different methods. The results are shown in the table with ** if the best algorithm is significantly better than the other methods. There are cases that the best method significantly differs from the others except for one/two method(s). In these cases, we indicate the method with *.

Discussion

We found that our proposed method, EDPCGRL, outperforms the baseline methods in the *Spiders Presented* metric, i.e., PCGRL showed fewer spiders to our virtual subjects before finding one that induces the specific subject's stress level on average. This might be because it combines exploitation and exploration. Figure 3 shows the *Spiders Presented* for each virtual subject in the sequence they were presented to all approaches. It reveals that the RL agent with the q-table initialized to 0 consistently outperforms other methods, which shows the method did not just learn our subjects' behaviour over time.

Our results reveal that for each stress level category: low, moderate, and high, it might be ideal to set the initial state attributes to the minimum, average, and maximum, respectively. However, we emphasize that increasing/decreasing all the spider's attributes does not always produce spiders closer to our desired one due to the variance in the virtual subject fitness functions. Nevertheless, in these situations, the Random baseline performs near optimally, because it is already closer to the desired spider in most cases.

We evaluate these methods on the virtual subjects, which simplifies the problem since the subjects are deterministic and do not change over time. However, if we imagine our virtual subjects not as distinct individuals, but as the same subject over time (e.g. fig. 3), we can observe that RL would also likely do better with dynamic, real-world individuals.

Initial State	Stress	Metric	Random	Greedy	GA	RL_Random	RL_Zero
Min	Low	Spiders Presented	$5.32\pm7.75^*$	11.24 ± 1.58	9.94 ± 1.76	6.17 ± 8.69	$\mathbf{4.04\pm2.08}^*$
		Accuracy	99.26	100	100	99.33	99.90
	Moderate	Spiders Presented	11.15 ± 5.24	21.95 ± 2.48	25.27 ± 5.66	$\mathbf{8.43\pm3.70}^{**}$	11.37 ± 4.84
		Accuracy	100	99	100	99.90	97.83
	High	Spiders Presented	39.86 ± 18.81	37.38 ± 4.14	49.09 ± 10.31	$\mathbf{33.10\pm17.72}^{**}$	29.04 ± 9.83
		Accuracy	81.90	63.33	100	89.76	53
Avg	Low	Spiders Presented	32.06 ± 19.66	27.38 ± 1.49	29.82 ± 5.97	$\mathbf{19.17\pm13.52}^{**}$	16.83 ± 9.17
		Accuracy	83.6	96	100	93.06	72.63
	Moderate	Spiders Presented	$5.11\pm6.63^*$	8.33 ± 3.23	10 ± 0	5.62 ± 5.39	$\mathbf{3.95\pm3.66}^*$
		Accuracy	100	100	100	100	99.06
	High	Spiders Presented	30.42 ± 20.17	$\mathbf{20.57\pm1.71}^{**}$	23.34 ± 5.78	23.50 ± 13.15	22.50 ± 11.00
		Accuracy	85.5	97.66	100	90.93	52.93
Max	Low	Spiders Presented	38.01 ± 18.67	37.40 ± 4.14	50.05 ± 10.31	$\mathbf{24.77\pm15.16}^{**}$	22.51 (8.49)
		Accuracy	81.30	63.66	100	93.06	70.93
	Moderate	Spiders Presented	11.45 ± 5.27	23.95 ± 2.48	25.24 ± 6.38	$9.29\pm4.49^*$	$\mathbf{8.56\pm3.32}^*$
		Accuracy	100	99	100	99.93	99.40
	High	Spiders Presented	$5.39\pm8.52^*$	11.24 ± 1.58	9.88 ± 1.72	$4.83\pm5.30^*$	$\mathbf{4.20\pm5.32}^*$
		Accuracy	99.43	100	100	99.20	95.90

Table 3: Mean and standard deviation for our Spiders Presented metric and the Accuracy of each approach across initial states and goal stress values.
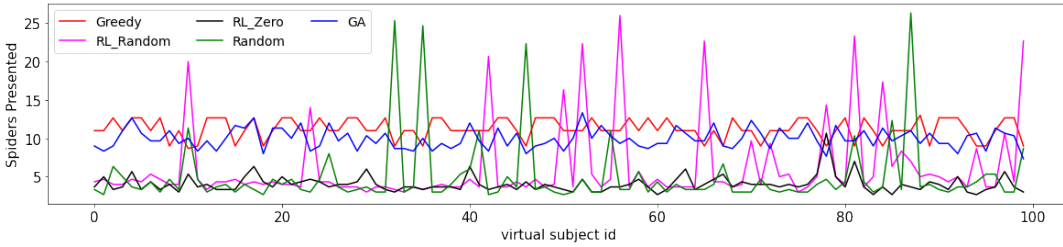

Figure 3: Spiders Presented for each virtual subject for the minimum initial state with target stress level set to low.

Therefore, our results show that our method has potential for real subjects.

We intend to evaluate the proposed method on real human subjects who are complicated and may respond differently over time. Therefore, it is challenging to obtain a model that maps physiological measures to stress levels. In fact, the same individuals may show different patterns in their physiological measures. Environmental factors may affect individuals' physiological response, e.g., drinking coffee, sleeping less/more. We require a robust and reliable model in our future work that accurately estimates stress levels from physiological measures. We also plan to improve the fidelity of the virtual spiders by using more sophisticated methods, such as more attributes and/or more possible ordinal values.

Currently, our approach uses a simple yet effective RL method. However, we are interested in investigating more sophisticated RL methods to determine if their performance differs significantly. There are many other algorithms available that utilize the combination of exploitation and exploration, e.g., Monte Carlo Tree Search (MCTS). However, they would require hyperparameter finetuning to work prop-

erly in our domain. We are also interested in applying transfer learning approaches to adapt the knowledge learned from one subject to a new subject.

Another limitation of our work is related to the ethical aspects. Our framework might cause excessive stress in the subjects. Therefore, the action selection policy in our PC-GRL should be carefully adjusted in a way that the RL agent considers gradually increasing the stress level. Nevertheless, the framework might also re-traumatize the subjects.

Conclusion

This paper introduces a new research area, i.e., EDPCGRL which was unidentified in prior work. We defined a proof of concept of our EDPCGRL framework for virtual reality exposure therapy, where PCGRL adjusts game content according to a subjects' physiological measures. We hypothesized that our EDPCG framework could be applied in computer-assisted therapy rather than reliant on pre-authored methods. We support this hypothesis by evaluating our proposed framework for arachnophobia in a case study. Different spiders with six different attributes are generated based on a

subject's stress level to find one that induces a target stress level. Our goal was to design a method that finds a desired spider with fewer spiders presented to a subject. We found that EDPCGRL outperformed existing experience-driven SBPCG methods for this task.

Acknowledgements

References

Abdessalem, H. B.; and Frasson, C. 2017. Real-time brain assessment for adaptive virtual reality game: A neurofeedback approach. In *international conference on brain function assessment in learning*, 133–143. Springer.

Badia, S. B.; Quintero, L. V.; Cameirao, M. S.; Chirico, A.; Triberti, S.; Cipresso, P.; and Gaggioli, A. 2018. Toward emotionally adaptive virtual reality for mental health applications. *IEEE journal of biomedical and health informatics* 23(5): 1877–1887.

Correa, O.; Cuervo, C.; Perez, P. C.; and Arias, A. 2014. A new approach for self adaptive video game for rehabilitation-experiences in the amblyopia treatment. In *2014 IEEE 3nd International Conference on Serious Games and Applications for Health (SeGAH)*, 1–5. IEEE.

Côté, S.; and Bouchard, S. 2009. Cognitive mechanisms underlying virtual reality exposure. *CyberPsychology & Behavior* 12(2): 121–129.

Cunningham, J.; Lopez, C.; Ashour, O.; and Tucker, C. S. 2020. Multi-Context Generation in Virtual Reality Environments Using Deep Reinforcement Learning. In *International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, volume 83983, V009T09A072. American Society of Mechanical Engineers.

Dimovska, D.; Jarnfelt, P.; Selvig, S.; and Yannakakis, G. N. 2010. Towards procedural level generation for rehabilitation. In *Proceedings of the 2010 Workshop on Procedural Content Generation in Games*, 1–4.

Earle, S.; Edwards, M.; Khalifa, A.; Bontrager, P.; and Togelius, J. 2021. Learning Controllable Content Generators. *arXiv :2105.02993* .

Freeman, D.; Haselton, P.; Freeman, J.; Spanlang, B.; Kishore, S.; Albery, E.; Denne, M.; Brown, P.; Slater, M.; and Nickless, A. 2018. Automated psychological therapy using immersive virtual reality for treatment of fear of heights: a single-blind, parallel-group, randomised controlled trial. *The Lancet Psychiatry* 5(8): 625–632.

Gisslén, L.; Eakins, A.; Gordillo, C.; Bergdahl, J.; and Tollmar, K. 2021. Adversarial Reinforcement Learning for Procedural Content Generation. *arXiv :2103.04847* .

Guzdial, M.; Liao, N.; and Riedl, M. 2018. Co-creative level design via machine learning. *arXiv :1809.09420* .

Heloir, A.; Nunnari, F.; Haudegond, S.; Havrez, C.; Lebrun, Y.; and Kolski, C. 2014. Design and evaluation of a self adaptive architecture for upper-limb rehabilitation. In *ICTs for Improving Patients Rehabilitation Research Techniques*, 196–209. Springer.

Hocine, N.; Gouaïch, A.; Cerri, S. A.; Mottet, D.; Froger, J.; and Laffont, I. 2015. Adaptation in serious games for upper-limb rehabilitation: an approach to improve training outcomes. *User Modeling and User-Adapted Interaction* 25(1): 65–98.

Khalifa, A.; Bontrager, P.; Earle, S.; and Togelius, J. 2020. Pcgrl: Procedural content generation via reinforcement learning. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 16, 95–101.

Kritikos, J.; Alevizopoulos, G.; and Koutsouris, D. 2021. Personalized Virtual Reality Human-Computer Interaction for Psychiatric and Neurological Illnesses: A Dynamically Adaptive Virtual Reality Environment That Changes According to Real-Time Feedback From Electrophysiological Signal Responses. *Frontiers in Human Neuroscience* 15.

Liapis, A.; Yannakakis, G. N.; and Togelius, J. 2013. Sentient sketchbook: computer-assisted game level authoring. In *8th International Conference on the Foundations of Digital Games (FDG 2013)*. ACM.

Lindner, P.; Miloff, A.; Reuterskiöld, L.; Andersson, G.; and Carlbring, P. 2019. What is so frightening about spiders? Self-rated and self-disclosed impact of different characteristics and associations with phobia symptoms. *Scandinavian journal of psychology* 60(1): 1–6.

Lopez, C. E.; Ashour, O.; and Tucker, C. S. 2019. Reinforcement learning content generation for virtual reality applications. In *ASME 2019 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*. American Society of Mechanical Engineers Digital Collection.

López, C. E.; Cunningham, J.; Ashour, O.; and Tucker, C. S. 2020. Deep Reinforcement Learning for Procedural Content Generation of 3D Virtual Environments. *Journal of Computing and Information Science in Engineering* 20(5).

Maples-Keller, J. L.; Bunnell, B. E.; Kim, S.-J.; and Rothbaum, B. O. 2017. The use of virtual reality technology in the treatment of anxiety and other psychiatric disorders. *Harvard review of psychiatry* 25(3): 103.

Miloff, A.; Lindner, P.; Dafgård, P.; Deak, S.; Garke, M.; Hamilton, W.; Heinsoo, J.; Kristoffersson, G.; Rafi, J.; Sindemark, K.; et al. 2019. Automated virtual reality exposure therapy for spider phobia vs. in-vivo one-session treatment: A randomized non-inferiority trial. *Behaviour research and therapy* 118: 130–140.

Nam, S.; and Ikeda, K. 2019. Generation of Diverse Stages in Turn-Based Role-Playing Game using Reinforcement Learning. In *2019 IEEE Conference on Games (CoG)*, 1–8. IEEE.

Risi, S.; Lehman, J.; D'Ambrosio, D. B.; Hall, R.; and Stanley, K. O. 2015. Petalz: Search-based procedural content generation for the casual gamer. *IEEE Transactions on Computational Intelligence and AI in Games* 8(3): 244–255.

Schmidt, P.; Reiss, A.; Duerichen, R.; Marberger, C.; and Van Laerhoven, K. 2018. Introducing wesad, a multimodal dataset for wearable stress and affect detection. In *Proceedings of the 20th ACM international conference on multimodal interaction*, 400–408.

Shaker, N. 2016. Intrinsically motivated reinforcement learning: A promising framework for procedural content generation. In *2016 IEEE Conference on Computational Intelligence and Games (CIG)*, 1–8. IEEE.

Shaker, N.; Togelius, J.; and Nelson, M. J. 2016. *Procedural content generation in games*. Springer.

Shiban, Y.; Schelhorn, I.; Pauli, P.; and Mühlberger, A. 2015. Effect of combined multiple contexts and multiple stimuli exposure in spider phobia: A randomized clinical trial in virtual reality. *Behaviour research and therapy* 71: 45–53.

Smith, A. M.; and Mateas, M. 2011. Answer set programming for procedural content generation: A design space approach. *IEEE Transactions on Computational Intelligence and AI in Games* 3(3): 187–200.

Summerville, A.; Snodgrass, S.; Guzdial, M.; Holmgård, C.; Hoover, A. K.; Isaksen, A.; Nealen, A.; and Togelius, J. 2018. Procedural content generation via machine learning (PCGML). *IEEE Transactions on Games* 10(3): 257–270.

Susanto, E. K.; and Tjandrasa, H. 2021. Applying Hindsight Experience Replay to Procedural Level Generation. In *2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)*, 427–432. IEEE.

Sutton, R. S.; and Barto, A. G. 2018. *Reinforcement learning: An introduction*. MIT press.

Togelius, J.; Yannakakis, G. N.; Stanley, K. O.; and Browne, C. 2011. Search-based procedural content generation: A taxonomy and survey. *IEEE Transactions on Computational Intelligence and AI in Games* 3(3): 172–186.

Vetter, R. S. 2013. Arachnophobic entomologists: when two more legs makes a big difference. *American Entomologist* 59(3): 168–175.

Yannakakis, G. N.; and Togelius, J. 2011. Experience-driven procedural content generation. *IEEE Transactions on Affective Computing* 2(3): 147–161.

Zahabi, M.; and Razak, A. M. A. 2020. Adaptive virtual reality-based training: a systematic literature review and framework. *Virtual Reality* 1–28.