

# IART – CHECKPOINT 2

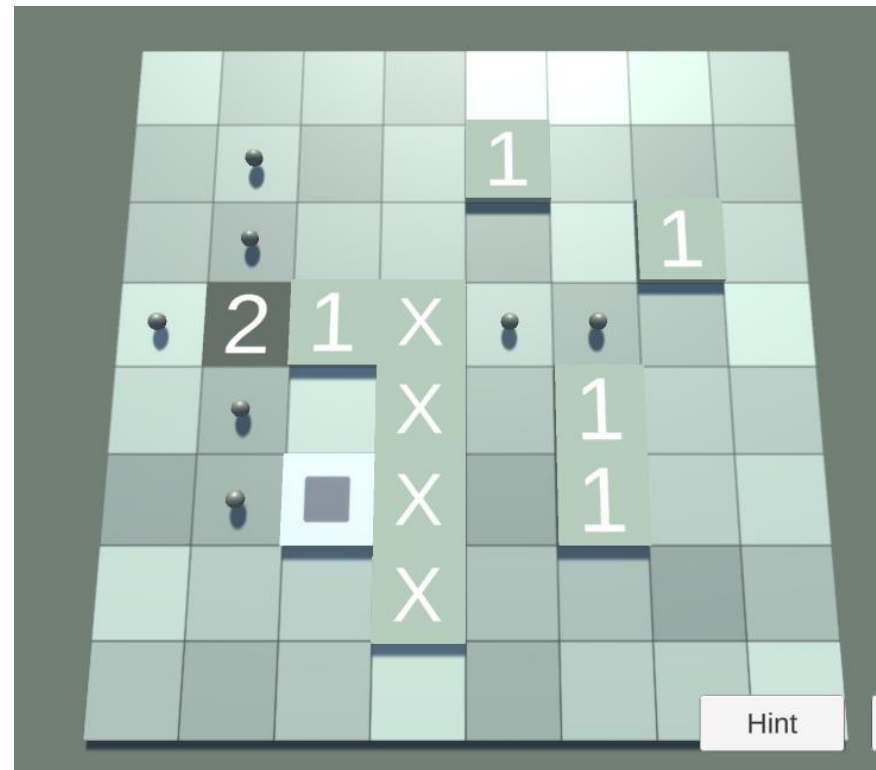
Daniel Brandão – 201705812

Gaspar Pinheiro – 201704700

Pedro Moás – 201705208

# Zhed – Aprendizagem por reforço

- Neste trabalho, vamos continuar a utilizar o jogo que trabalhamos no primeiro projeto, Zhed, desenvolvido em unity e c#. Desta vez, utilizaremos algoritmos de aprendizagem por reforço para ajudar o computador a resolver os diferentes puzzles, através da ferramenta ML-agents do Unity.
- Para isso, é possível que tenhamos de, ao início, reduzir a complexidade do jogo ou dos níveis a solucionar.



# Trabalho Relacionado

- Deep Reinforcement Learning for 2048 <http://www.mit.edu/~amarj/files/2048.pdf>
- Improving Generalization Ability in a Puzzle Game Using Reinforcement Learning [http://www.cig2017.com/wp-content/uploads/2017/08/paper\\_71.pdf](http://www.cig2017.com/wp-content/uploads/2017/08/paper_71.pdf)
- Applying Reinforcement learning to find the logic puzzles solution <https://journals.tuplovdiv.bg/index.php/journal/article/view/16/14>
- Deep Reinforcement Learning Bot That Plays Tetris <https://github.com/nuno-faria/tetris-ai>
- Playing Tetris with Deep Reinforcement Learning [http://cs231n.stanford.edu/reports/2016/pdfs/121\\_Report.pdf](http://cs231n.stanford.edu/reports/2016/pdfs/121_Report.pdf)
- ML Agents Learning Environment Examples – GridWorld <https://github.com/Unity-Technologies/ml-agents/blob/master/docs/Learning-Environment-Examples.md#gridworld>

# Ferramentas

Iremos utilizar ML-agents, plugin Unity de código aberto que permite que jogos e simulações sirvam como ambientes para treinar agentes inteligentes, e C# para continuar a utilizar o projeto anterior como base.

# Algoritmos

- Tendo em conta que estamos limitados às capacidades do plugin ML-agents, apenas poderemos utilizar os algoritmos que este tem disponibilizado, Proximal Policy Optimization (PPO) e Soft Actor Critic (SAC).
- O algoritmo SAC difere do PPO na medida em que é um algoritmo off-policy, sendo que o computador poderá aprender das experiências obtidas em qualquer episódio anterior.
- Iremos comparar os diferentes algoritmos, assim como variar os parâmetros entre experiências sobre o mesmo algoritmo, de modo a analisar qual o método mais eficaz.

# Desenvolvimento até agora

- Implementação do jogo em Unity e c#
- Setup do ambiente de treino
- Experimentação com múltiplos exemplos de ML-agents
- Adaptação e experimentação do Zhed no ML-agents.

