Jogando RTS com Aprendizado de Reforço Profundo

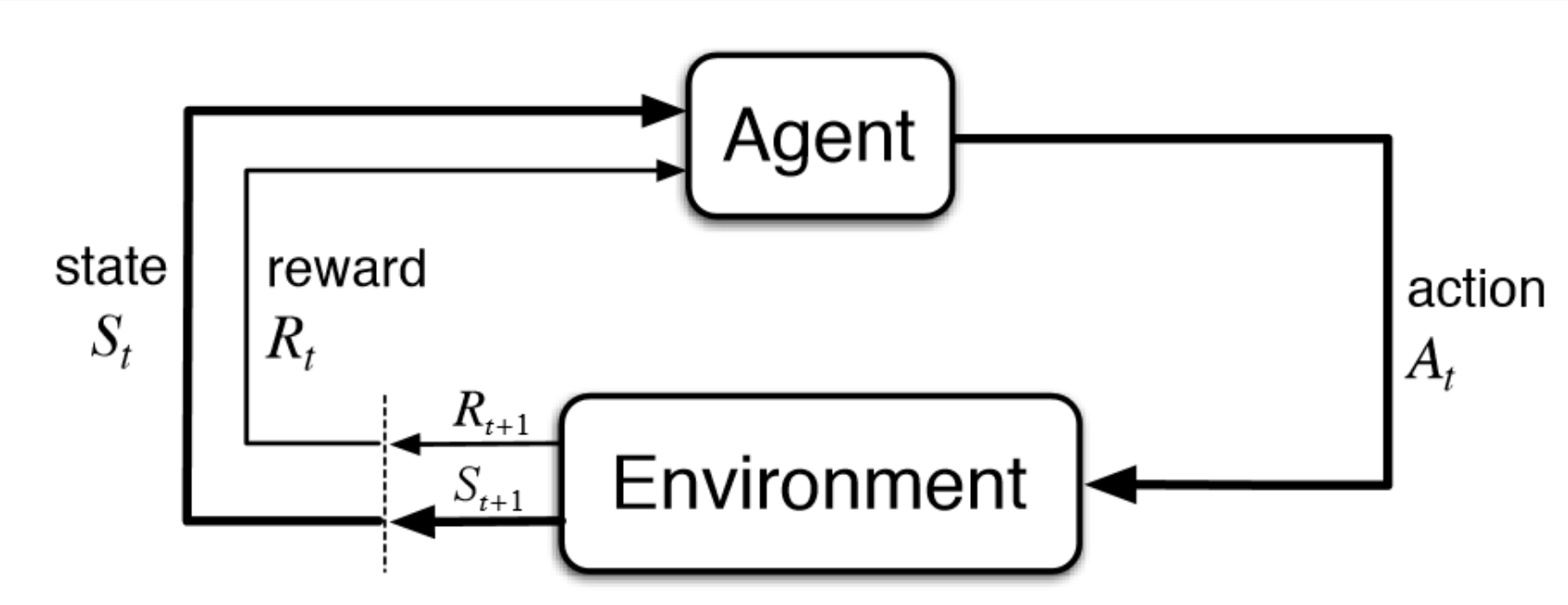
Victor Aliende da Matta
Supervisor: Prof. Dr. Denis Deratani Mauá

DCC - Instituto de Matemática e Estatística da USP

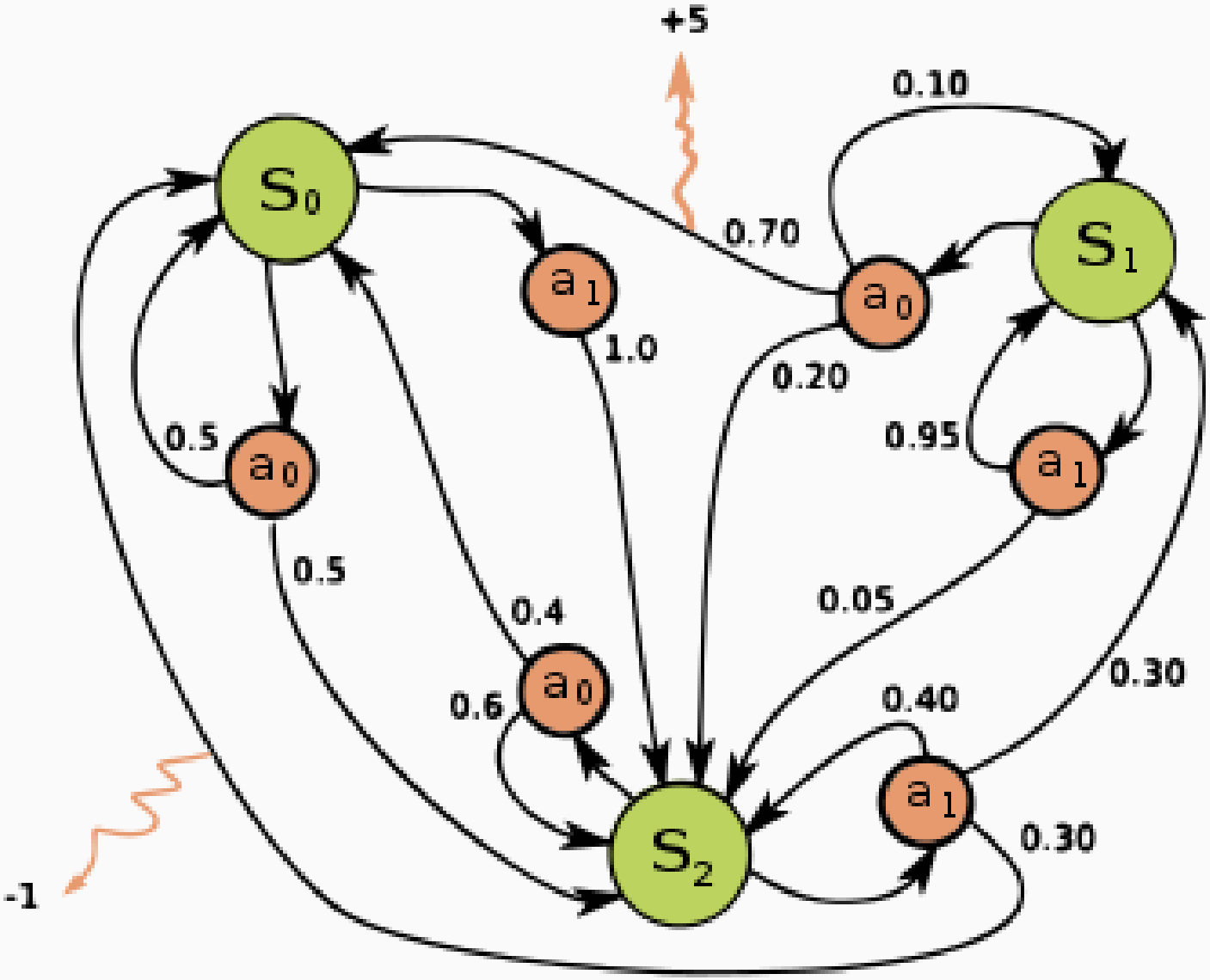
Sumário

Aprendizado de Reforço

Agente - Ambiente - Recompensa



Processo de Decisão de Markov



Definições

- Política $\pi(a|s)$
- Episódio, Retorno descontado
- Função valor $v_\pi(s)$

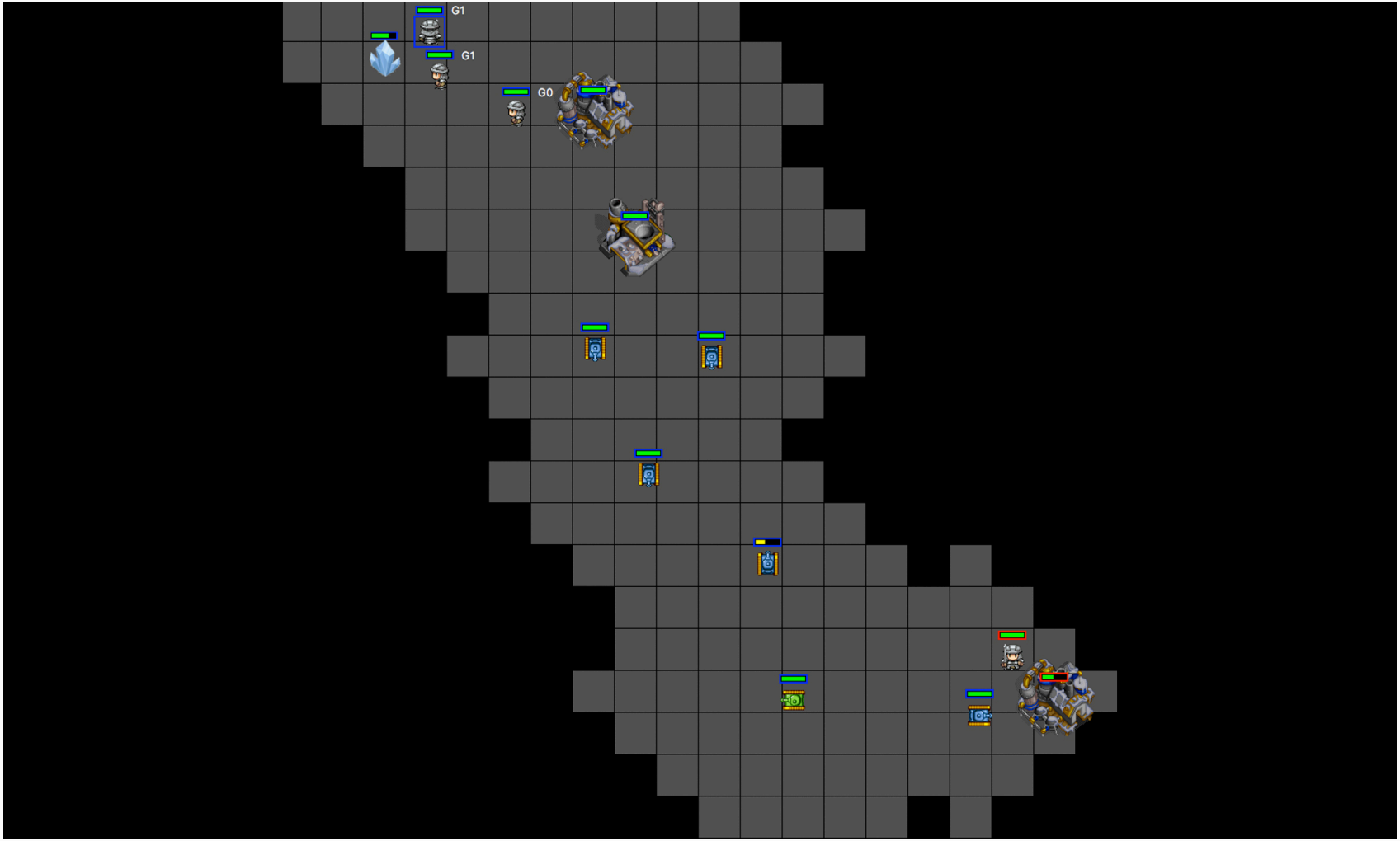
O Ambiente

RTS

Porque jogar RTS?

- Desafio complexo e bem definido
- Ambiente totalmente controlável
- Conjunto de estados grande
- Informação imperfeita
- Recompensas esparsas e distantes

MiniRTS



Algoritmos Actor-Critic

Gradiente de Política

Abordagem Geral

1. Parametrização da política $\pi(a|s, \theta)$
2. Definição da função objetivo: $\mathcal{J}(\theta) = v_{\pi_\theta}(S_0)$
3. Subida de gradiente

Teorema do Gradiente de Política

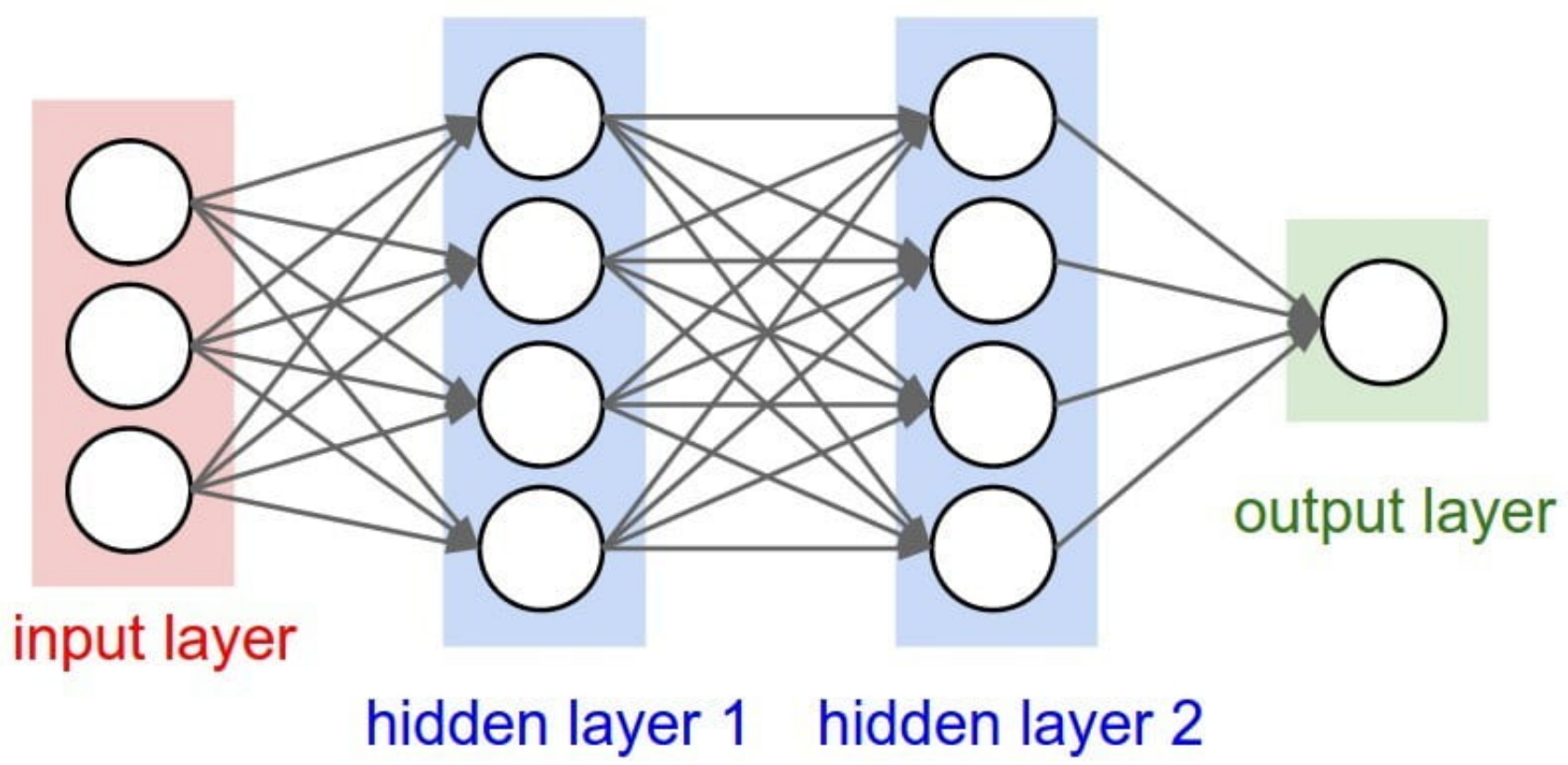
$$\nabla \mathcal{J}(\theta) \propto \mathbb{E}_{\pi}\left[q_{\pi}(S_t, A_t)\frac{\nabla_{\theta}\pi_{\theta}(A_t|S_t)}{\pi_{\theta}(A_t|S_t)}\right] = \mathbb{E}_{\pi}\left[G_t\frac{\nabla_{\theta}\pi_{\theta}(A_t|S_t)}{\pi_{\theta}(A_t|S_t)}\right]$$

Algoritmos

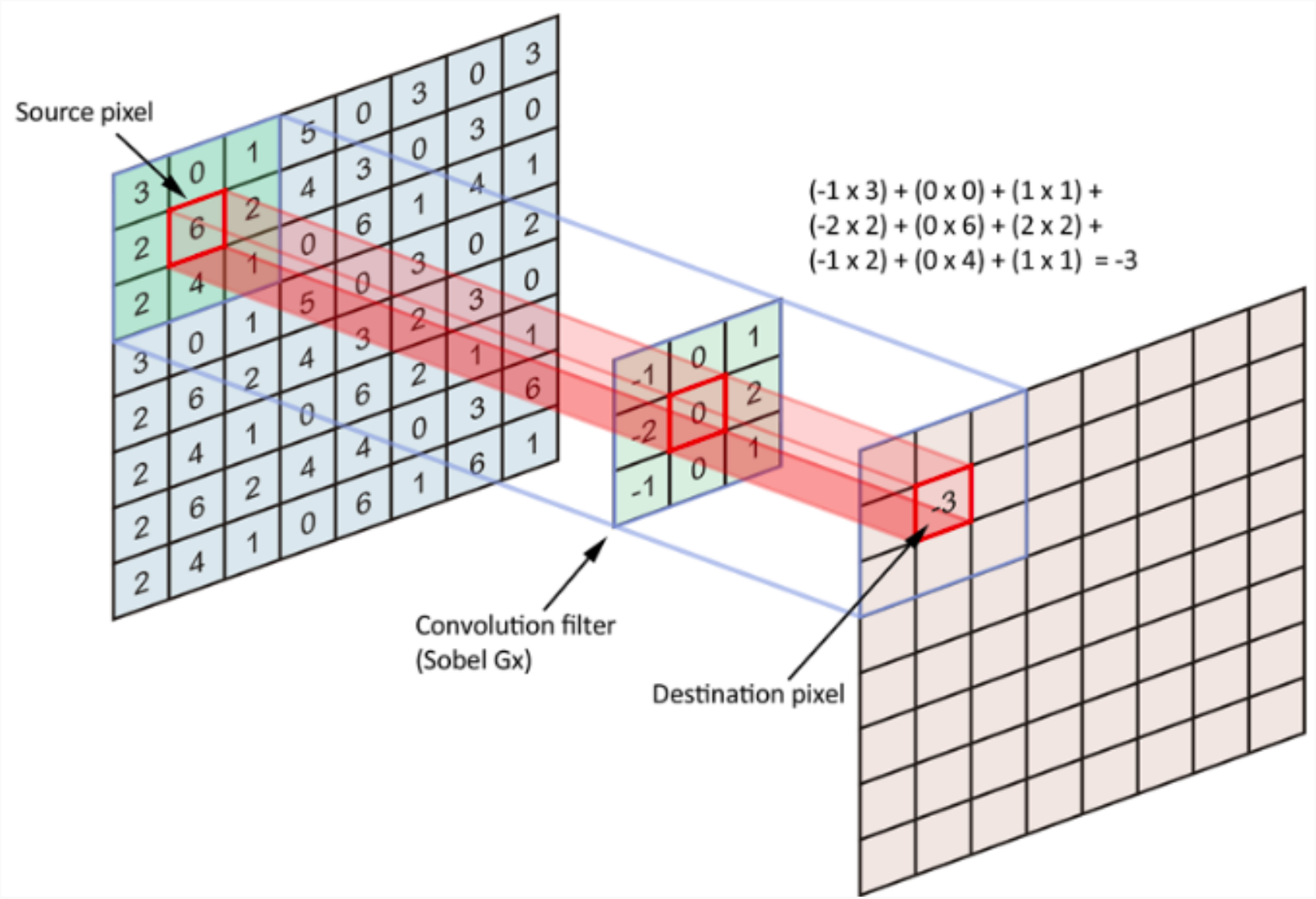
- REINFORCE
- Actor-Critic
- A3C - *Asynchronous Advantage Actor-Critic*

O Modelo

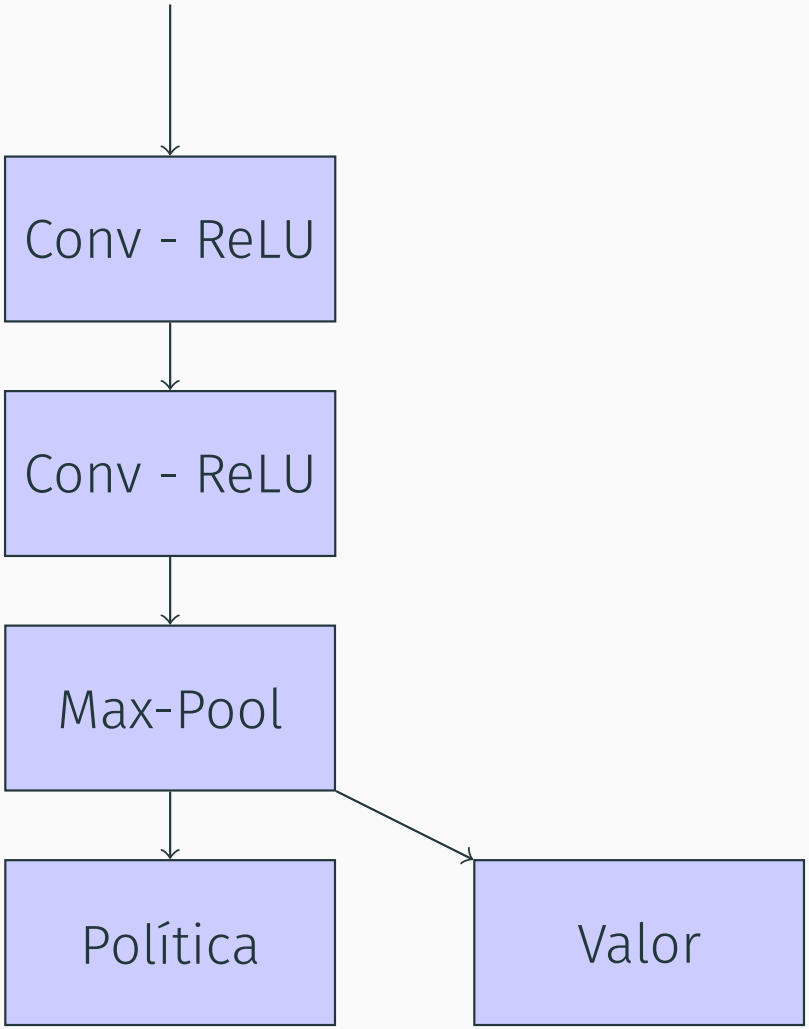
Redes Neurais



Convolução


Source pixel
(-1 x 3) + (0 x 0) + (1 x 1) +
(-2 x 2) + (0 x 6) + (2 x 2) +
(-1 x 2) + (0 x 4) + (1 x 1) = -3
-3
Convolution filter
(Sobel Gx)
Destination pixel

Nossa Arquitetura



Resultados

Resultados

Abordagem	Acurácia
Agente aleatório vs AI_SIMPLE	0.259
Agente aletório vs AI_HIT_AND_RUN	0.242
vs AI_SIMPLE	0.537
vs AI_HIT_AND_RUN	0.671
Tian et al. vs AI_SIMPLE	0.684
Tian et al. vs AI_HIT_AND_RUN	0.636

O FIM

References i


V. Mnih, A. P. Badia, M. Mirza, A. Graves, T. P. Lillicrap, T. Harley, D. Silver, and K. Kavukcuoglu.
Asynchronous methods for deep reinforcement learning.
CoRR, abs/1602.01783, 2016.


R. S. Sutton and A. G. Barto.
Reinforcement learning: An introduction.
2011.


Y. Tian, Q. Gong, W. Shang, Y. Wu, and C. L. Zitnick.
Elf: An extensive, lightweight and flexible research platform for real-time strategy games.
In *Advances in Neural Information Processing Systems*, pages 2659–2669, 2017.