

An aerial view of a racetrack with green grass and blue runoff areas, set against a backdrop of a city and a body of water. Two red sports cars are shown in the upper left, appearing to fly or race over the track.

# AI Racers

Trabalho prático 2 da disciplina de IA

Felipe Colombelli  
Giovanna Miotto  
Henrique Silva

# Implementação



## ◆ Estrutura de dados da tabela Q

- ◇ Dicionário de dicionários
- ◇ Chaves: estados
- ◇ Valores: dicionário
  - ◇ Chaves: ações
  - ◇ Valores: valor Q daquela ação naquele estado

# Exploração



## ◆ Boltzmann

- ◇ Temperatura inicial: 90
- ◇ Arrefecimento linear de 0.995
- ◇ Bound final: 2
- ◇ Estagnação: episódio 762

## ◆ Epsilon Greedy

- ◇ Epsilon inicial: 0.1 (praticamente random sempre)
- ◇ Aumenta linearmente sendo dividido por um coeficiente de 0.99
- ◇ Bound final: 0.98
- ◇ Estagnação: episódio 228

# Features / Estado



## ◆ **Primeiro conjunto** (utilizado na competição)

### ◆ Estado e níveis de discretização:

- ◇ speed - 5
- ◇ dist\_ahead - 5
- ◇ dist\_left - 3
- ◇ dist\_rigth - 3
- ◇ dist\_bomb - 2
- ◇ angle\_bomb - 3
- ◇ on\_ice - 2
- ◇ on\_grass - 2

# Reward



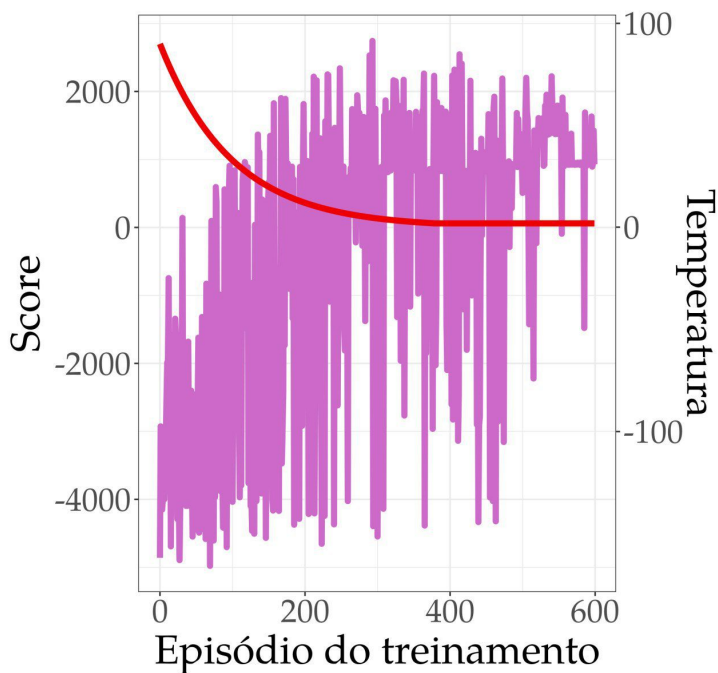
## ◆ **Primeiro conjunto** (utilizado na competição)

- ◆ Se está na pista (20 | -20)
- ◆ Diff (ou 20 caso passou um checkpoint)
- ◆ Bomb warning
- ◆ Speed bonus (ganha 5 se ação foi acelerar)

# Resultados primeiro conjunto



Resultados do baseline



# Features / Estado



## ◆ Segundo conjunto

◆ Estado e níveis de discretização:

- ◇ Mesma coisa que do primeiro conjunto com a adição de:
- ◇ `check_diff - 5`

# Reward



- ◆ R\_ON\_TRACK = 10
- ◆ P\_NOT\_ON\_TRACK = -15
- ◆ R\_BACK\_ON\_TRACK = 10
- ◆ P\_WAYWARD\_DRIVER = -10
  - ◇ \*6 se estava na pista acelerou e entrou na grama
  - ◇ \*2 se freiou e tentou evitar
  - ◇ \*1 se tentou dobrar
  - ◇ \*4 se não fez nada
  - ◇ \*6 se estava na grama e freiou ou não fez nada
- ◆ URGE\_CHECKPOINT = -10
- ◆ BOMB\_WARNING = -15



# Reward - continuação

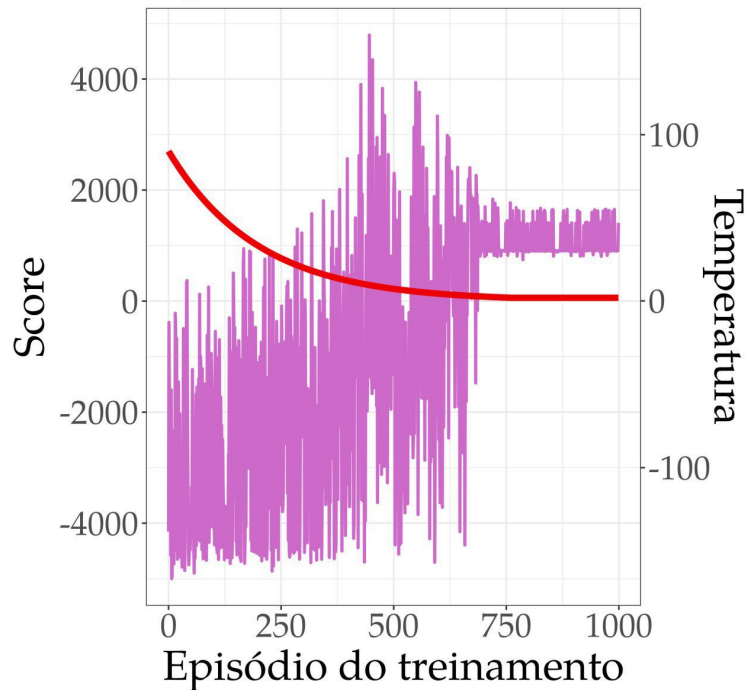


- ◆  $\text{diff} * 2$  (ou 40 se passou por um checkpoint)
- ◆  $\text{velocidade} / 10$

# Resultados segundo conjunto



Updated -  $T = 90$



# Terceiro conjunto

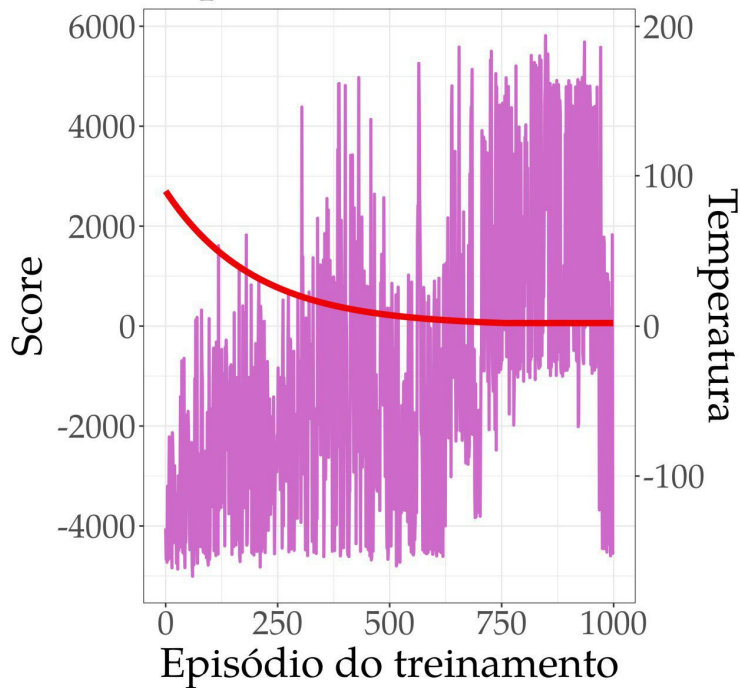


- ◆ Uma versão com speed boost para o conjunto anterior
- ◆ Mudança: reward
  - ◇  $\text{velocidade} / 10 \Rightarrow \text{velocidade}$

# Resultados terceiro conjunto



Super -  $T = 90$



# Quarto conjunto

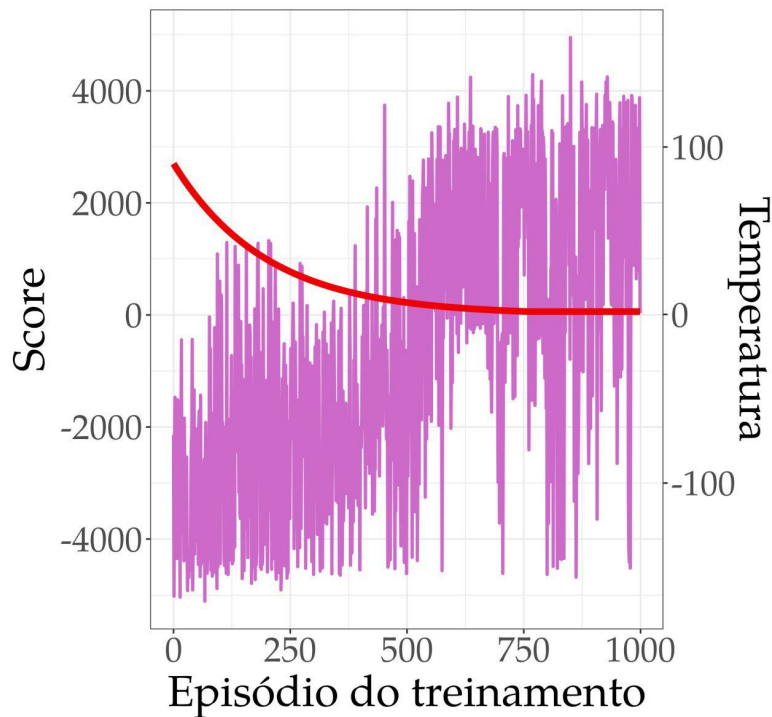


- ◆ Uma versão com speed boost moderado para o conjunto anterior
- ◆ Melhores resultados
- ◆ **Mudança:** reward
  - ◇ velocidade  $\Rightarrow$  velocidade / 5

# Resultados quarto conjunto



Moderado -  $T = 90$



# Resultados quarto conjunto



Vídeo de demonstração...

# Quarto conjunto

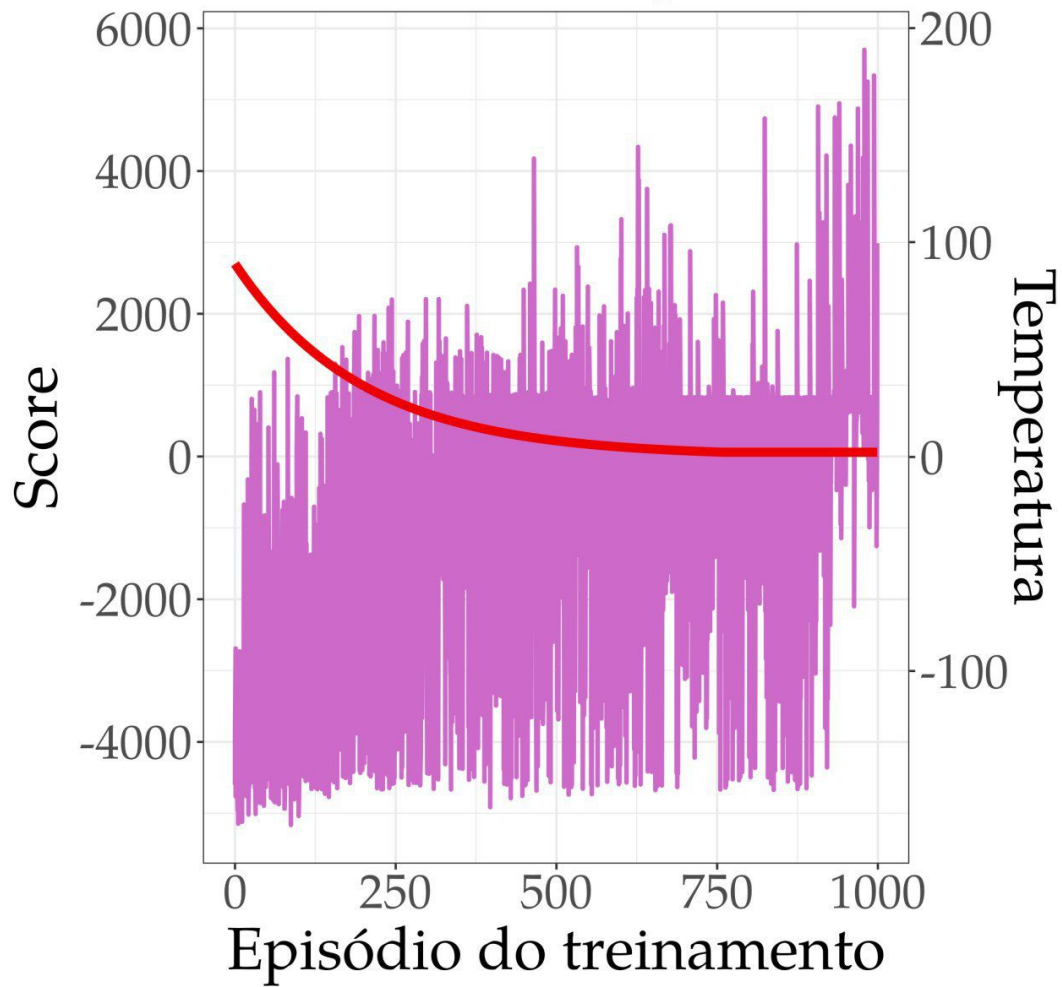


## ◆ Variações:

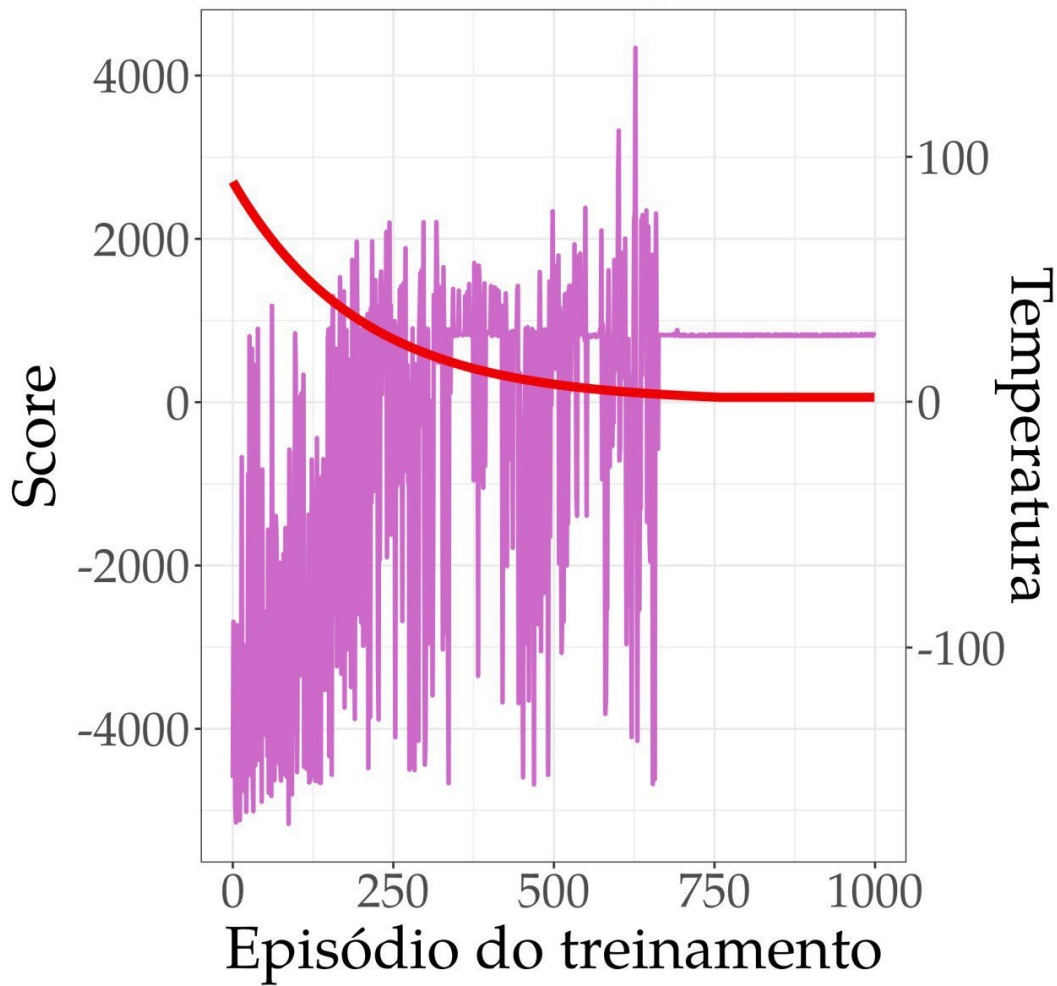
- ◇ Alpha
- ◇ Gamma
- ◇ Temperatura inicial
- ◇ Estratégia



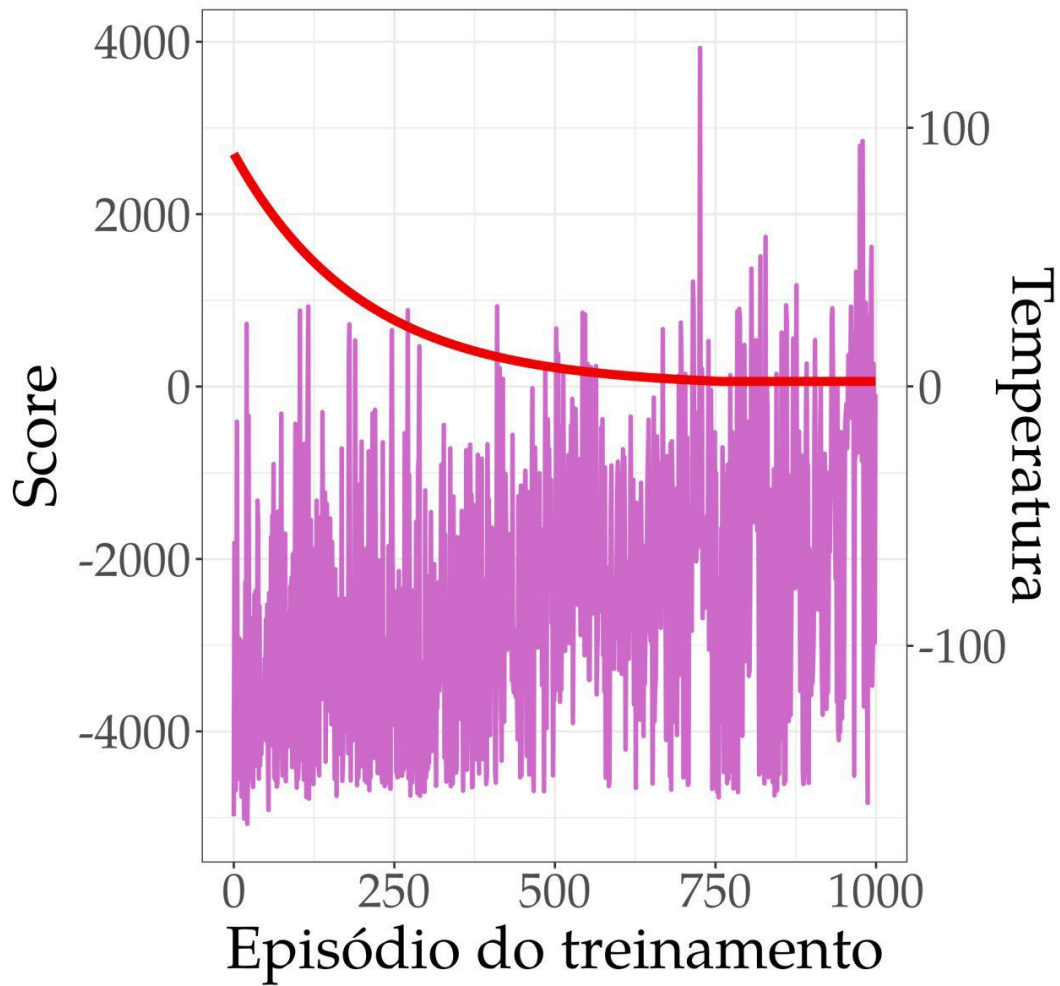
# Moderado - $\alpha = 0.3$



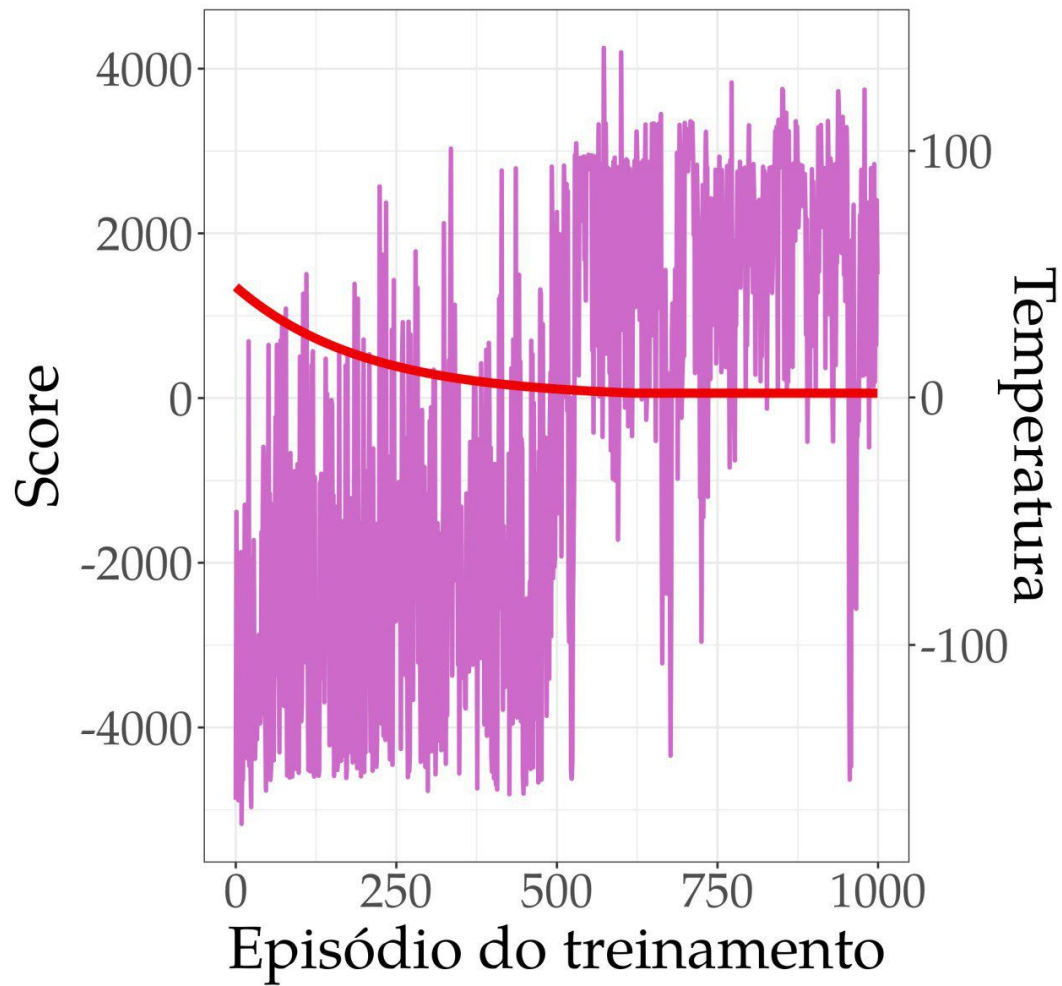
# Moderado - $\alpha = 0.8$



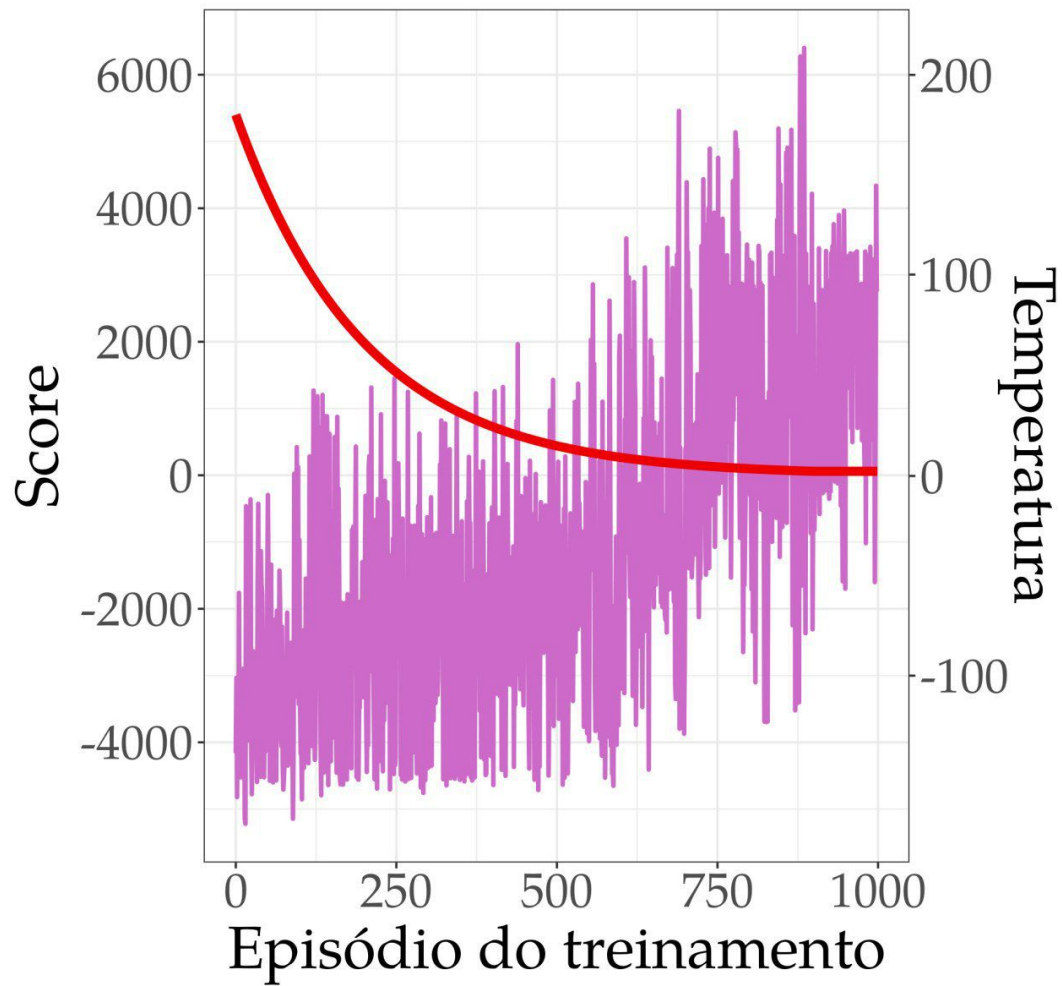
# Moderado - $\gamma = 0.8$



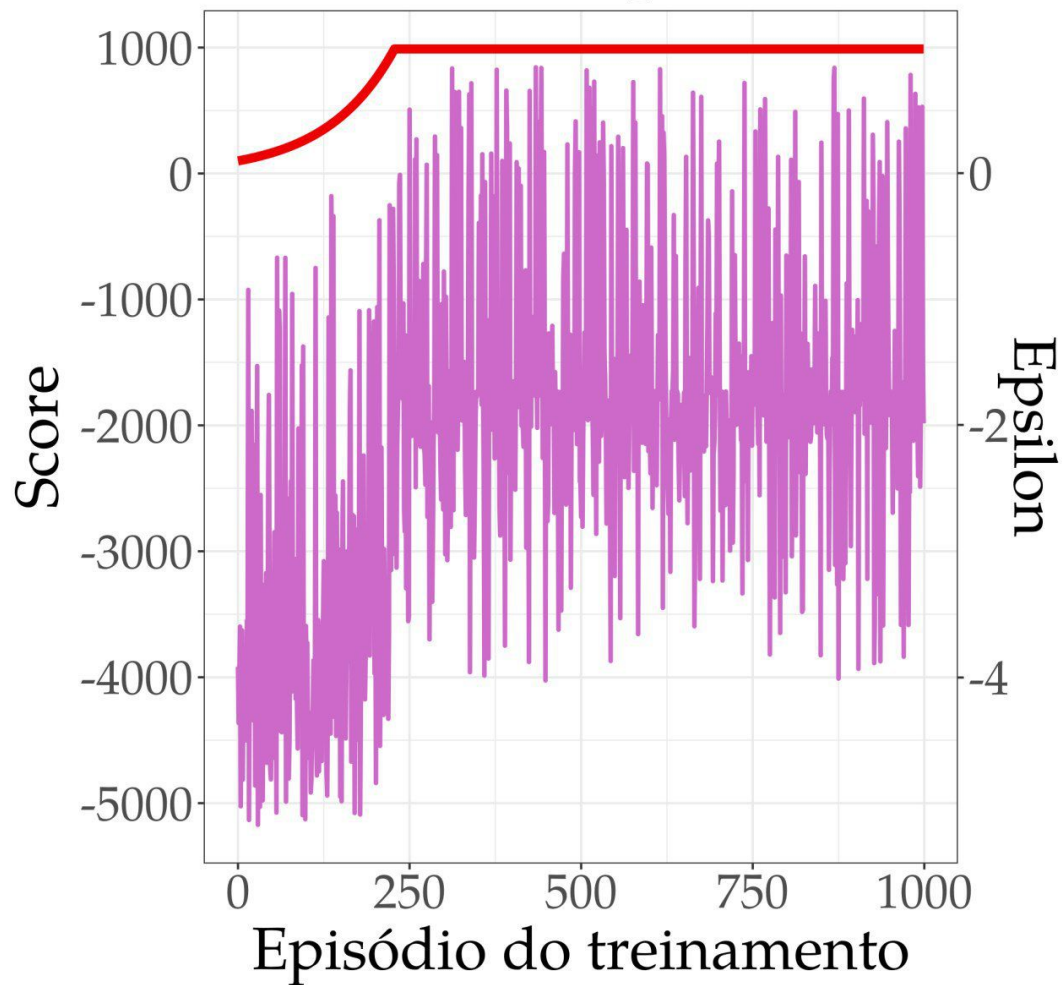
# Moderado - $T = 45$



# Moderado - $T = 180$



# Moderado - epsilon



# Conclusão



- ◆ Boltzmann se mostrou bem mais eficiente
- ◆ Mais iterações deveriam ser feitas para tentar atingir uma convergência
- ◆ Um speed boost foi decisivo na melhora do agente
- ◆ Diferentes níveis de discretizações deveriam ser testados