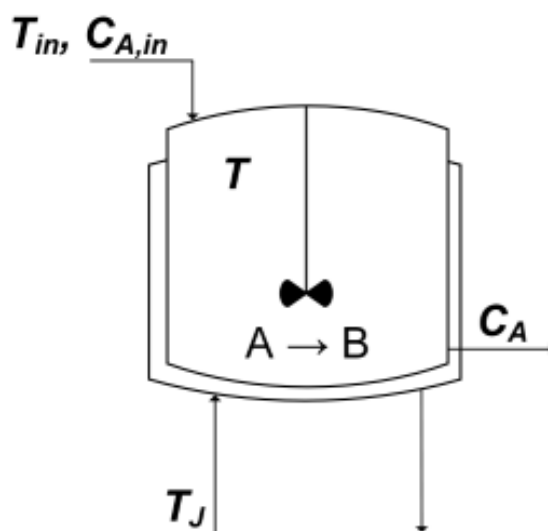


Trabalho de Reinforcement Learning (Prazo para entrega: 11/12/2024)

Considere o reator CSTR representado pela figura abaixo, no qual ocorre a reação exotérmica desejada $A \rightarrow B$.



Deseja-se utilizar o algoritmo de aprendizado por reforço DDPG (Lillicrap et al., 2016), que é um algoritmo do tipo Actor-Critic, para desenvolver um agente que controle a concentração de A na saída do reator, manipulando a temperatura do fluido na camisa.

A descrição do ambiente e do algoritmo usado para o treino do controlador estão disponíveis <https://arxiv.org/pdf/2410.22093> e suas implementações no notebook Jupyter:

https://colab.research.google.com/drive/19AJ2q_43-9IRBG1fT0ztULh12OUtjL_l?usp=sharing

O objetivo do trabalho é treinar o agente para otimizar a resposta do sistema para variações no valor de C_A^{SP} .

- 1) Substitua a função recompensa padrão usada no notebook (r_squared) pela função recompensa abaixo:

$$R(t) = \begin{cases} +1, & \text{se } (1 - a)C_A^{SP} < C_A < (1 + a)C_A^{SP} \\ 0, & \text{senão} \end{cases}; \text{ onde } a \in]0,0.2]$$

- a) Para $a = 0.1$, quais são o retorno médio e a resposta de C_A média obtidos no cenário de teste?
- b) A partir de uma análise de sensibilidade sobre a e o número de passos de treinos, melhore os resultados obtidos para o cenário de teste. Então, para o melhor resultado obtido, mostre a resposta de C_A conseguida no cenário de teste e o retorno médio conseguido.

- 2) Implemente a sua própria função recompensa e verifique o retorno médio e a resposta de C_A média obtidos nesse cenário.

O trabalho deve ser encaminhado via e-mail a bruno@eq.ufrj.br e joseneto@eq.ufrj.br antes de 11/12/2024, o assunto do e-mail sendo “Trabalho RL EQE778”, com os seguintes arquivos:

- um relatório em formato pdf com as respostas às questões
- os códigos das funções recompensa implementadas

Referências

TP Lillicrap, JJ Hunt, A Pritzel, N Heess, T Erez, Y Tassa, D Silver, and D Wierstra. Continuous Control with deep reinforcement learning. ICRL conference paper, 2016