
Instituto Tecnológico de Costa Rica**Escuela de Ingeniería Electrónica****Trabajo Final de Graduación****Proyecto:** Método basado en aprendizaje reforzado para el control automático de una planta no lineal.**Estudiante:** Oscar Andrés Rojas Fonseca

I Semestre 2024

Firma del asesor

Bitácora de trabajo

Fecha	Actividad	Anotaciones	Horas dedicadas
19/03/2024	1. Continuación de pruebas para la definición de las clases referentes a <i>CartPole</i> .	a) La definición de las clases se complica dadas las condiciones de los ejemplos lineales en [1] para el <i>MPC</i> . b) Se continua con la definición en específico de la clase <i>CartPoleSystem</i> .	6 horas
20/03/2024	2. Reunión de seguimiento con el asesor del proyecto.	a) Revisión de avance en el código y errores de forma. b) Dada la complejidad de la adaptación de los ambientes <i>CartPole</i> al <i>MPC</i> [1], se decidió revisar otras fuentes, encontrando un ejemplo de implementación <i>CartPole</i> con <i>DQN</i> [2].	2 horas
20/03/2024	3. Pruebas de montaje del código <i>DQN</i> mencionado.	a) Creación de nuevo <i>env</i> para el funcionamiento del código en [2]. b) Pruebas de funcionamiento del código original (exitosa). c) Pruebas de implementación del código con su versión CUDA para mejor desempeño en SO Windows.	4 horas

21/03/2024	4. Pruebas con entorno <i>CartPole</i> para implementación de CUDA.	a) Se hicieron pruebas de funcionamiento con CUDA en Windows resultado erróneo.	5 horas
21/03/2024	5. Revisión del código original para <i>DQN</i> [2].	a) El código referencia se encuentra bien ordenado y comentado en secciones importantes, muestra potencial.	4 horas
Total de horas de trabajo:			21 horas

Contenidos de actividades

Luego de las pruebas fallidas con el *MPC* de Airaldi [1], con consejo del profesor asesor se optó por buscar otro método para el control, encontrando un ejemplo mediante *DQN* [2] aplicado al mismo entorno *CartPole*, primeros pasos para la adaptación al PAMH, de manera que se probó el código y el resultado fue exitoso y con clara reducción de episodios de entrenamiento (de 600 a 50) para implementación sin CUDA, ejemplo en Figura 1

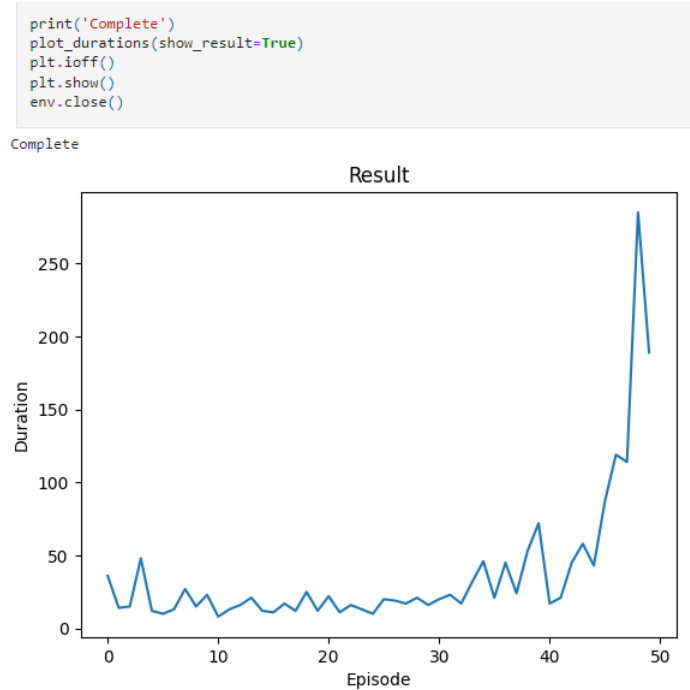


Figure 1: Resultado de entrenamiento del modelo para *CartPole* con 50 episodios.

Referencias

- [1] F. Airaldi, A. Bietti, A. Casagrande, and A. Bemporad, “Learning model predictive control with policy gradients,” *IEEE Transactions on Automatic Control*, 2023.
- [2] A. Paszke and M. Towers, “Reinforcement learning (dqn) tutorial,” *PyTorch*.