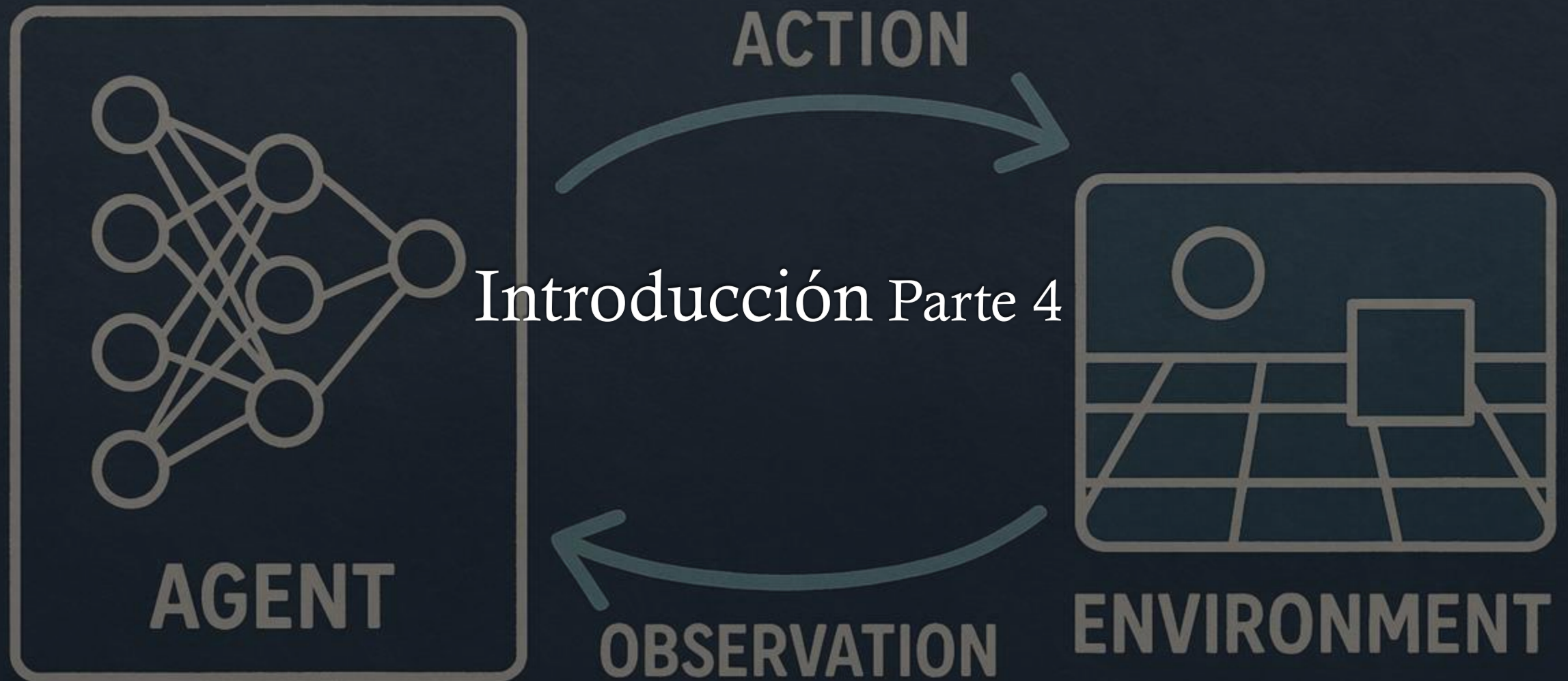


DEEP REINFORCEMENT LEARNING



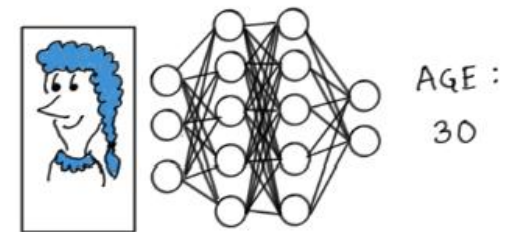
Identificar tareas para RL

Si respondes sí a cualquiera de estas preguntas, el aprendizaje por refuerzo (RL) es una buena opción

1. ¿Es difícil saber cuál es la acción correcta de antemano (difícil ser un maestro), pero más fácil evaluar una acción o secuencia de acciones después de que se ha realizado (fácil dar recompensas)?
2. ¿Los datos dependen de datos y decisiones anteriores?
3. ¿La tarea requiere *pensamiento a largo plazo, estrategia o coordinación*?



Train a bot to trade stocks



Predict age from images

¿Cuándo es mejor usar RL?

- ◇ Cuando las decisiones son secuenciales
- ◇ Las acciones conducen a la exploración del entorno.

Tomemos como ejemplo la robótica, una aplicación clásica del RL.

- ◇ El objetivo de un robot es aprender a realizar tareas desconocidas.
- ◇ No deberías indicarle cómo lograrlo, ya sea porque es demasiado difícil (¿cómo le dices a un robot que haga una pirueta?)
- ◇ o porque tu propia experiencia (no eres un robot) podría influir en tu comportamiento, por lo que desconoces la mejor manera de moverte como tal.
- ◇ Si, en cambio, permites que el robot explore, puede iterar hacia una solución óptima.

Quiz

1 ¿Cuál es el objetivo en un problema de RL?

- ◊ Maximizar el número de acciones.
- ◊ Maximizar la sumatoria de rewards.
- ◊ Maximizar la precision de las acciones.

Quiz

2 ¿Cuáles de las siguientes indica que un problema es apto para RL?

- ◊ Hay mucho datos anotados disponibles
- ◊ Las observaciones en cualquier punto dependen de observaciones pasadas o decisions previas.
- ◊ El problema requiere planeación a largo término
- ◊ El problema requiere habilidades de percepción (visión, habla, escucha)
- ◊ Es difícil enseñar por ejemplos pero es fácil saber si la acciones tomadas son buenas

Referencias

- [1] Hu, M. (2023). *The Art of Reinforcement Learning*. Springer Science and Business Media LLC.
- [2] Stein, H. (2023) Introduction to Reinforcement Learning.
- [3] <https://medium.datadriveninvestor.com/supervised-and-unsupervised-learning-7281050992a0>
- [4] <https://www.sharpsightlabs.com/blog/regression-vs-classification/>
- [5] <https://nixustechnologies.com/unsupervised-machine-learning/>
- [6] <https://www.bombaysoftwares.com/blog/introduction-to-unsupervised-learning>
- [7] He, K., Chen, X., Xie, S., Li, Y., Dollár, P., & Girshick, R. (2022). Masked autoencoders are scalable vision learners. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 16000-16009).
- [8] 6.S191 Introduction to Deep Learning introtodeeplearning.com
- [9] Barto, A.G.; Bradtke, S.J.; Singh, S.P. Learning to act using real-time dynamic programming. *Artif. Intell.* **1995**, *72*, 81–138.
- [10] Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *nature*, 518(7540), 529-533.

Referencias

- [11] <https://deepmind.google/discover/blog/deepmind-ai-reduces-google-data-centre-cooling-bill-by-40/>
- [12] <https://www.ft.com/content/16b8ffb6-7161-11e7-aca6-c6bd07df1a3c>
- [13] <https://www.wired.com/story/game-makers-inject-ai-develop-more-lifelike-characters/>
- [14] Walk, Run, Crawl, RL Fun | Boston Dynamics | Atlas
https://www.youtube.com/watch?v=I44_zbEwz_w
- [15] Stepping Up | Reinforcement Learning with Spot | Boston Dynamics
<https://www.youtube.com/watch?v=Kf9WDqYKYQQ>
- [16] Fast Deep RL Course. Dibya's School. <https://courses.dibya.online/>
- [17] Winder, P. (2020). *Reinforcement learning*. O'Reilly Media.
- [18] <https://amturing.acm.org/byyear.cfm>
- [19] June 2025 CACM: Richard S. Sutton and Andrew G. Barto, Recipients of ACM's 2024 A.M. Turing Award, <https://www.youtube.com/watch?v=RrXibq7-W6o>