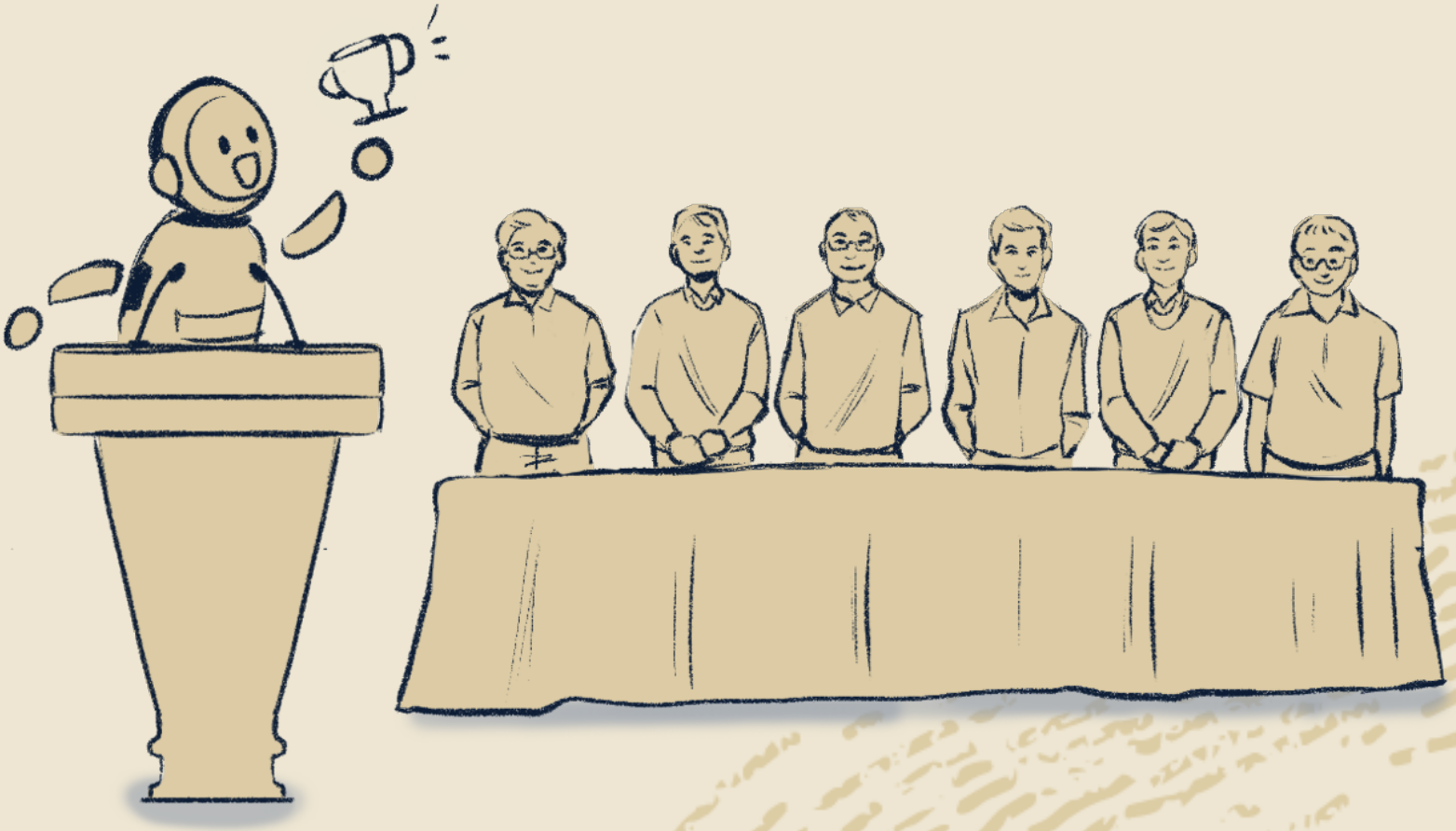
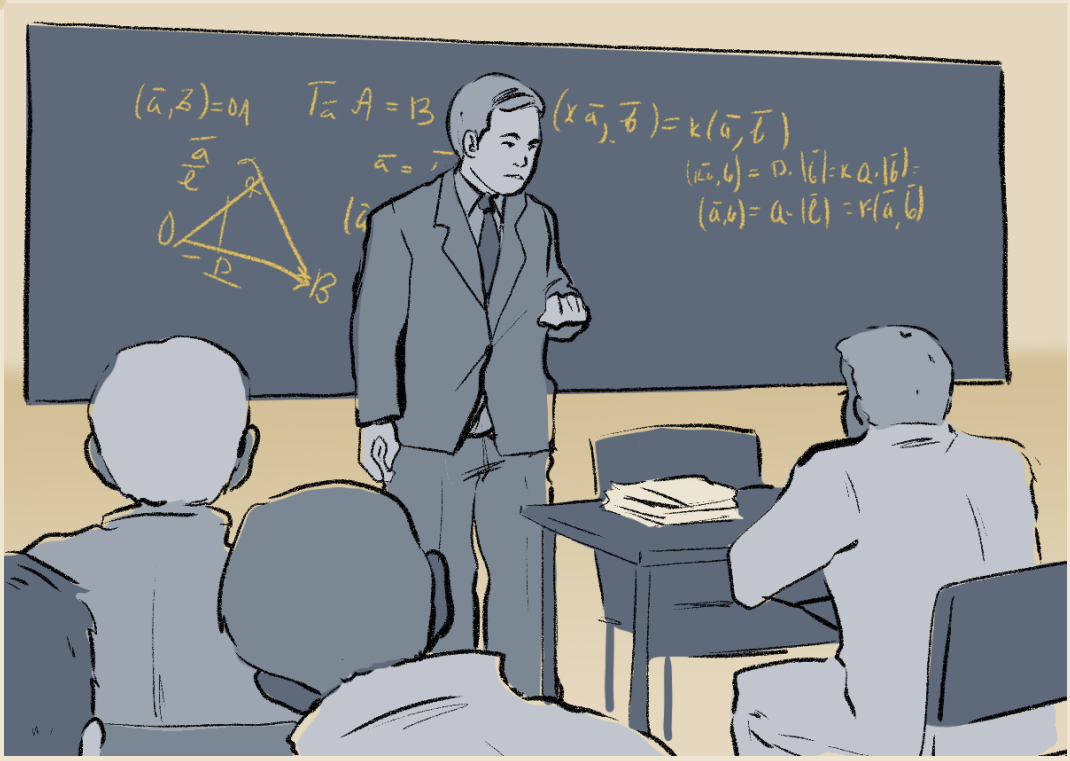
BOURBAKI

COLEGIO DE MATEMÁTICAS

Deep Reinforcement Learning en 25 artículos for the working scientist





Esta selección de **25 artículos** ha sido creada como acompañamiento al curso *Deep Reinforcement Learning*, una formación intensiva que combina los fundamentos del aprendizaje por refuerzo con las capacidades expresivas del deep learning. A diferencia de los enfoques tradicionales de aprendizaje supervisado, aquí los agentes aprenden *actuando*, *equivocándose* y *mejorando* progresivamente a través de la interacción con su entorno.

En esta curaduría se abordan los algoritmos más importantes del DRL, desde los métodos clásicos como Q-Learning y la ecuación de Bellman, hasta variantes modernas y robustas como PPO, GRPO y Conservative Q-Learning (CQL), incluyendo casos de uso reales que ilustran su impacto en la industria y la investigación.

1. Reinforcement Learning: An Introduction. MIT Press.

2. Human-level control through deep reinforcement learning. Nature, 518(7540), 529–533.

3. Deep Reinforcement Learning with Double Q-learning. AAAI.

4. Proximal Policy Optimization Algorithms. arXiv preprint.

5. Trust Region Policy Optimization. ICML.

6. Continuous control with deep reinforcement learning. ICLR.

7. Soft Actor-Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor. ICML.

8. Addressing Function Approximation Error in Actor-Critic Methods. ICML.

9. Conservative Q-Learning for Offline Reinforcement Learning. NeurIPS.

10. Policy invariance under reward transformations: Theory and application to reward shaping. ICML.

11. Deterministic Policy Gradient Algorithms. ICML.

12. OpenAI Gym. arXiv preprint.

13. Rainbow: Combining Improvements in Deep Reinforcement Learning. AAAI.

14. Unifying count-based exploration and intrinsic motivation. NeurIPS.

15. Hindsight Experience Replay. NeurIPS.

16. Prioritized Experience Replay. ICLR.

17. Mastering the game of Go with deep neural networks and tree search. Nature.

18. Action-conditional video prediction using deep networks in Atari games. NeurIPS.

19.

Maximum entropy inverse reinforcement learning. AAAI.

20.

End-to-end training of deep visuomotor policies. JMLR.

21.

Learning to reach goals via iterated forward search. ICLR.

22.

DeepMimic: Example-guided deep reinforcement learning of physics-based character skills. ACM TOG.

23.

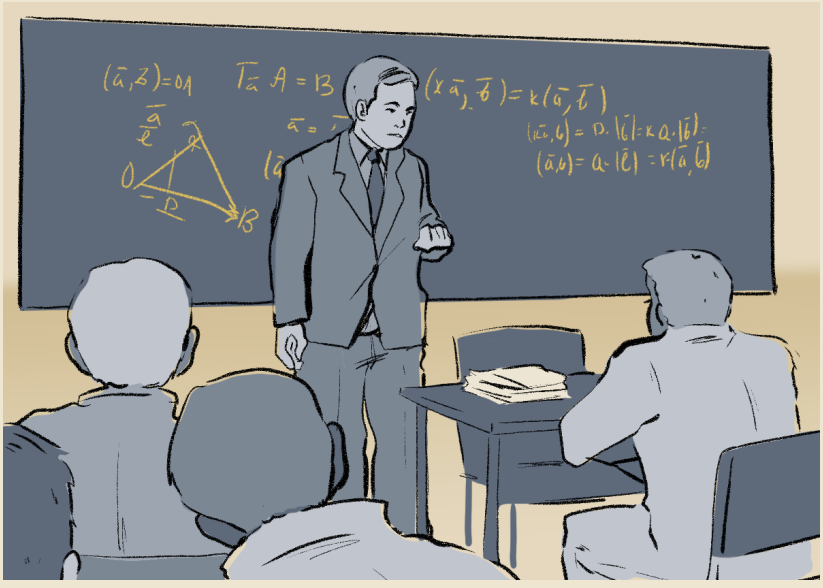
DeepMind Control Suite. arXiv preprint.

24.

RL with Human Feedback: Benchmarking RLHF algorithms. NeurIPS.

25.

Stable-Baselines3: Reliable Reinforcement Learning Implementations. JMLR.



BOURBAKI

COLEGIO DE MATEMÁTICAS

Estaremos encantados de explicar con
detalle sobre nuestros cursos

Escríbenos por mail o WhatsApp

info@colegio-bourbaki.com

+52 56 2141 7850