

## Aprendizaje por Refuerzos: Teoría y Aplicaciones en Robótica, Psicología y Neurociencias

Programa:

- Introducción a la toma de decisiones: los banditos de k-brazos. El dilema exploración-explotación.
  - Teorías de aprendizaje animal: Pavlov y el condicionamiento instrumental. Exploración-explotación en ratas (y humanos). Son sub-óptimas las ratas?
  - Problemas con las teorías de condicionamiento: el problema de los delayed rewards y la asignación de crédito.
  - Aprendizaje por refuerzos: cómo solucionar los problemas con las teorías de condicionamiento.
  - Formalizando AR: los Procesos de Decisión de Markov (PDM). Definición de valor. Resolviendo un PDM mediante programación dinámica y programación lineal.
  - Algoritmos de aprendizaje model-free: Q-learning, TD( $\lambda$ ), SARSA. El problema de la exploración en esquemas model-free.
  - Algoritmos de aprendizaje model-based: la familia R-Max. Exploración quasi-óptima en algoritmos model-based. El problema de planning.
  - Otras representaciones: PDM factorizados, relacionales y orientados a objetos.
  - Aplicaciones en robótica y videojuegos.
  - Aprendizaje por refuerzos en neurociencias: el rol de la dopamina en el cerebro.
  - Algoritmos de AR en el cerebro: estudios de electrofisiología y resonancia magnética funcional.
- Acciones de Documento**
- 

### 12.- BIBLIOGRAFÍA (indicar título del libro, autor, Editorial y año de publicación)(adjuntada)

- Abbeel, P., Coates, A., Quigley, M., & Ng, A. Y. (2007). An application of reinforcement learning to aerobatic helicopter flight. In Advances in Neural Information Processing Systems 19 (p. 2007). MIT Press.
- Barto, A. G., Sutton, R. S., & Watkins, C. J. C. H. (1989). Learning and sequential decision making. LEARNING AND COMPUTATIONAL NEUROSCIENCE (pp. 539-602). MIT Press.
- Bellman, R. (1957). Dynamic programming. Princeton University Press.
- 7
- Bertsekas, D. P., & Tsitsiklis, J. N. (1996). Neuro-dynamic programming. Belmont, MA: Athena Scientific.
- Crites, R., & Barto, A. (1996). Improving elevator performance using reinforcement learning. Advances in Neural Information Processing Systems 8 (pp. 1017-1023). MIT Press.
- Graepel, T., Herbrich, R., & Gold, J. (2004). Learning to fight. Proceedings of the International Conference on Computer Games: Artificial Intelligence, Design and Education.



Guez, A., Vincent, R. D., Avoli, M., & Pineau, J. (2008). Adaptive treatment of epilepsy via batch-mode reinforcement learning. AAAI (pp. 1671-1678). AAAI Press.

Ipek, E., Mutlu, O., Mart'ínez, J. F., & Caruana, R. (2008). Self-optimizing memory controllers: A reinforcement learning approach. ISCA '08: Proceedings of the 35th International Symposium on Computer Architecture (pp. 39-50). Washington, DC, USA: IEEE Computer Society.

Kaelbling, L. P., Littman, M. L., & Moore, A. P. (1996). Reinforcement learning: A survey. Journal of Artificial Intelligence Research, 4, 237-285.

Kohl, N., & Stone, P. (2004). Machine learning for fast quadrupedal locomotion. The Nineteenth National Conference on Artificial Intelligence (pp. 611-616).

Lee, H., Shen, Y., Yu, C.-H., Singh, G., & Ng, A. Y. (2006). Quadruped robot obstacle negotiation via reinforcement learning. ICRA (pp. 3003-3010). IEEE.

Merrick, K., & Maher, M. L. (2006). Motivated reinforcement learning for non-player characters in persistent computer game worlds. ACE '06: Proceedings of the 2006 ACM SIGCHI International conference on Advances in Computer Entertainment Technology (p. 3). New York, NY, USA: ACM.

Ng, A. Y., Kim, H. J., Jordan, M. I., & Sastry, S. (2004). Autonomous helicopter flight via reinforcement learning. Advances in Neural Information Processing Systems 16 (NIPS-03).

Puterman, M. L. (1994). Markov Decision Processes—Discrete Stochastic Dynamic Programming. New York, NY: John Wiley & Sons, Inc.

8

Silver, D., Sutton, R., & M"uller, M. (2007). Reinforcement learning of local shape in the game of Go. IJCAI.

Sutton, R. S. (1988). Learning to predict by the methods of temporal differences. MACHINE LEARNING (pp. 9-44). Kluwer Academic Publishers.

Sutton, R. S., & Barto, A. G. (1998). Reinforcement Learning: An Introduction. The MIT Press.

Tesauro, G. (1994). TD-Gammon, a self-teaching backgammon program, achieves master-level play. Neural Computation, 6, 215-219.

Tesauro, G. (2005). Online resource allocation using decompositional reinforcement learning. Proceedings, The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference, July 9-13, 2005, Pittsburgh, Pennsylvania, USA (pp. 886-891). AAAI Press / The MIT Press.

Tsitsiklis, J. N., & Roy, B. V. (2000). Regression methods for pricing complex american-style options. IEEE Transactions on Neural Networks, 12, 694-703.





Universidad de Buenos Aires  
Facultad de Ciencias Exactas y Naturales

Referencia Expte. N° 500.324/2011

Buenos Aires,

14 NOV 2011

**VISTO:**

la nota presentada por el Dr. Sebastián Uchitel, Director del Departamento de Computación, mediante la cual eleva la información y el programa del curso de posgrado **Aprendizaje por refuerzos: teoría y aplicaciones en robótica, psicología y neurociencias**, dictado en el segundo cuatrimestre de 2010 dictado por el Dr. Gregorio Diuk Wasser,

el CV del Dr. Gregorio Diuk Wasser

**CONSIDERANDO:**

que las presentes actuaciones no fueron presentados en tiempo y forma por un error administrativo,

lo actuado por la Comisión de Doctorado de esta Facultad el 25/10/2011,

lo actuado por la Comisión de Enseñanza, Programas, Planes de estudio y Posgrado

lo actuado por este Cuerpo en la sesión realizada en el día de la fecha,

en uso de las atribuciones que le confiere el Artículo 113° del Estatuto Universitario,

**EL CONSEJO DIRECTIVO DE LA FACULTAD DE  
CIENCIAS EXACTAS Y NATURALES  
RESUELVE:**

**Artículo 1°:** Dar validez al dictado del curso de posgrado **Aprendizaje por refuerzos: teoría y aplicaciones en robótica, psicología y neurociencias** de 15 hs. de duración, en el segundo cuatrimestre de 2010.

**Artículo 2°:** Aprobar el programa del curso de posgrado **Aprendizaje por refuerzos: teoría y aplicaciones en robótica, psicología y neurociencias**, obrante a fs 3 y 4 del expediente de referencia.

**Artículo 3°:** Aprobar un puntaje máximo para la Carrera de Doctorado de 0,5 (cero coma cinco) puntos

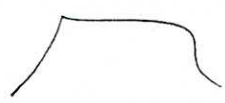
**Artículo 4°:** Aprobar un arancel de 20 Módulos. Disponer que los montos recaudados serán utilizados conforme a lo dispuesto por Resolución CD N° 072/03.

**Artículo 5°:** Comuníquese al Director del Departamento de Computación, a la Biblioteca de la FCEyN, a la Biblioteca de la FCEyN y a la Subsecretaría de Postgrado (con fotocopia del Programa incluido)

**Artículo 6°:** Comuníquese al Departamento de Alumnos (sin fotocopia del programa incluido). Cumplido Archívese.

Resolución CD N° 2871==  
SP/ga/25/10/2011

  
Dr. JAVIER LÓPEZ DE CASENAVE  
SECRETARIO ACADEMICO

  
Dr. JORGE ALIAGA  
DECANO