



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Planilla a completar para presentación de Cursos de Posgrado

1.- DEPARTAMENTO de COMPUTACIÓN

2.- NOMBRE DEL CURSO: Aprendizaje por Refuerzos

3.- DOCENTES:

RESPONSABLE/S: **Carlos Diuk**

COLABORADORES:

AUXILIARES:

4.- CARRERA de DOCTORADO

5.- AÑO: **2012**

CUATRIMESTRE/S: **1º Cuatrimestre**

6.- PUNTAJE PROPUESTO PARA CARRERA DE DOCTORADO: **2 puntos**

7.- DURACIÓN (anual, cuatrimestral, bimestral u otra): **bimestral**

8.- CARGA HORARIA SEMANAL:

Teóricas:

Problemas:

Laboratorio:

Seminarios:

Teórico – Práctico: **40 horas**

Salida a Campo:

9.- CARGA HORARIA TOTAL: **40 horas**

10.- FORMA DE EVALUACIÓN: parcial en la computadora, trabajo práctico, examen final

11.- PROGRAMA ANALÍTICO:

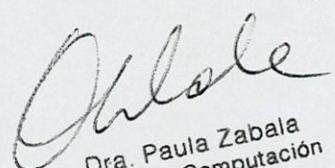
- Introducción a la toma de decisiones: definición del problema, antecedentes en psicología desde Pavlov hasta hoy, teorías de aprendizaje animal.



- Bandidos de k-brazos. El dilema exploración-explotación. Algoritmos (cuasi-)óptimos. Complejidad computacional vs. complejidad de aprendizaje.
- El problema de los delayed rewards y la asignación de crédito. Aprendizaje por refuerzos: cómo solucionar los problemas de las teorías de condicionamiento.
- Formalizando RL: los Procesos de Decisión de Markov (MDPs). Definición de valor. Algoritmos para resolver un MDP 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. Complejidad.
- Algoritmos de aprendizaje model-based: la familia R-Max. Exploración cuasi-óptima en algoritmos model-based. Complejidad. El problema de planning.
- Planning avanzado: Sample-based planning. Monte Carlo Tree Search y su impacto en el juego de Go.
- Representaciones avanzadas: MDP factorizados, relacionales y orientados a objetos. Algoritmos de aprendizaje bajo estas representaciones.
- Modelos jerárquicos: descomponiendo un problema en sub-problemas. Descubrimiento automático de sub-objetivos. Algoritmos para RL jerárquico.
- Aplicaciones en robótica. Breve introducción a algoritmos de policy-search, y descripción detallada de la solución utilizada en el vuelo de helicóptero invertido.
- Aplicaciones en juegos: backgammon, Go. Aplicaciones en videojuegos comerciales.
- Aprendizaje por refuerzos en neurociencias: el rol de la dopamina en el cerebro. Algoritmos de RL en el cerebro: estudios de electrofisiología y resonancia magnética funcional.
- Tópicos avanzados: se dedicarán las últimas 2 o 3 clases a otros tópicos avanzados, a decidir en base a los intereses del curso.

12.- BIBLIOGRAFÍA:

- Reinforcement Learning: An Introduction, Richard S Sutton y Andrew G Barto, MIT Press, Cambridge, MA, 1998. Disponible en: <http://www.incompleteideas.net/sutton/book/the-book.html>
- Algorithms for Reinforcement Learning, Csaba Szepesvari, Morgan & Claypool, 2010. Disponible en: <http://www.ualberta.ca/~szepesva/RLBook.html>
- Reinforcement Learning: A survey, Leslie Kaelbling, Michael Littman y Andrew Moore, Journal of Artificial Intelligence Research, 1996. Disponible en: <http://arxiv.org/pdf/cs.AI/9605103>
- Reinforcement learning: the good, the bad and the ugly, Peter Dayan y Yael Niv, Current Opinion in Neurobiology, 2008. Disponible en: <http://www.gatsby.ucl.ac.uk/~dayan/papers/dayanniv08.pdf>


Dra. Paula Zabala
Depto. de Computación
F.C.E.N. - U.B.A.



Universidad de Buenos Aires
Facultad de Ciencias Exactas y Naturales

Referencia Expte. N° 500.769/2012

Buenos Aires, 04 JUN 2012

VISTO:

la nota presentada por la Dra. Paula Zabala del Departamento de computación, mediante la cual eleva la información y el programa del curso de posgrado **Aprendizaje por Refuerzos**, dictado durante el primer cuatrimestre de 2012 por Carlos Diuk,

CONSIDERANDO:

lo actuado por la Comisión de Doctorado de esta Facultad el 02/05/2012
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°: Autorizar el dictado del curso de posgrado **Aprendizaje por Refuerzos** de 40 hs. de duración.

Artículo 2°: Aprobar el programa del curso de posgrado **Aprendizaje por Refuerzos**, obrante a fs 2 y 3 del expediente de referencia.

Artículo 3°: Aprobar un puntaje máximo de dos (2) puntos para la Carrera del Doctorado.

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 a la Dirección del Departamento de Computación, a la Biblioteca de la FCEyN y a la Subsecretaría de Postgrado (con fotocopia del Programa incluido) y a la Dirección de Alumnos y Graduados sin fotocopia del Programa. Cumplido Archívese.

Resolución CD N° 1102
SP/med/ 14/05/2012