

# Aprendizaje por refuerzo

Clase 25: Hitchhiker's guide to RL





# Antes de empezar...

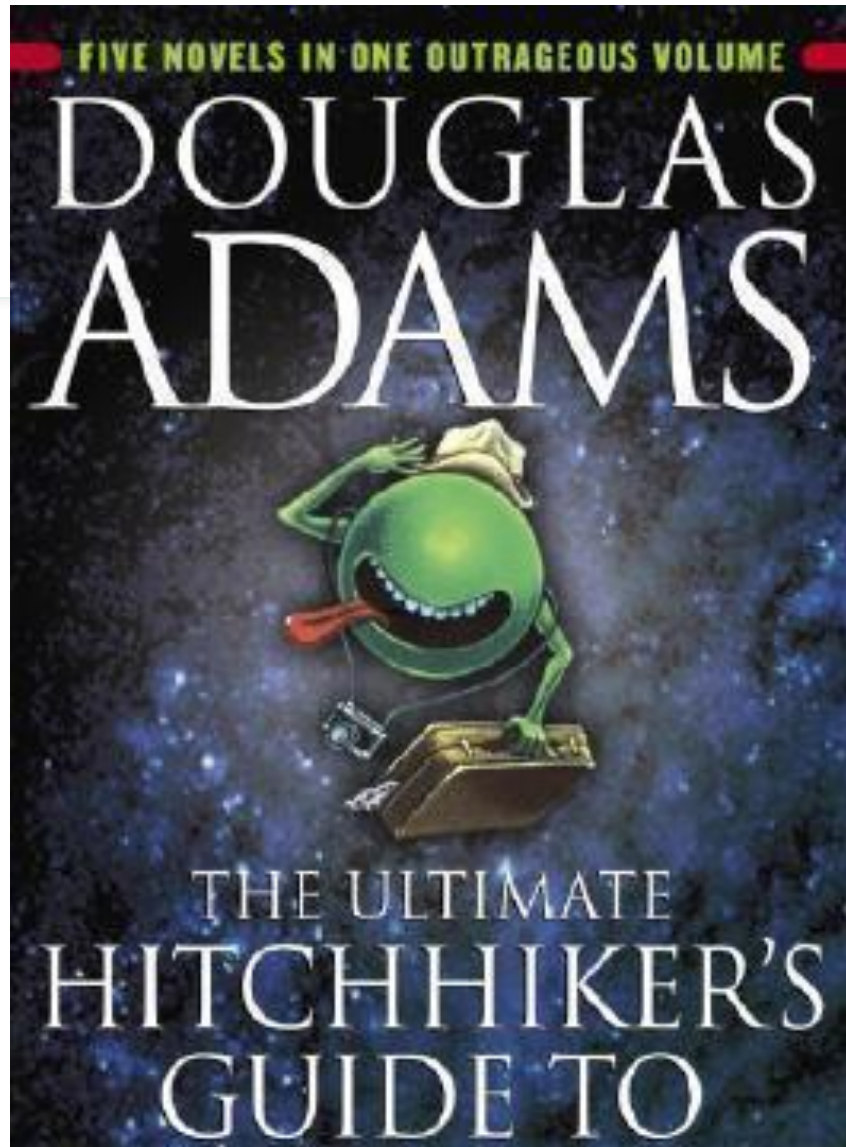
- Dudas de
  - Examen 22 de mayo
  - Proyecto 24 de mayo
  - Entrega tareas 26 de mayo
  - Finales 29 de junio



## Para el día de hoy...

- Lo que vimos
- Problemas abiertos
- Áreas de interés
- Publicando en RL
- Oportunidades





Aprendizaje  
por refuerzo



# Lo que saben (o deberían saber)...

- Algoritmos de programación dinámica
- Algoritmos con funciones de aproximación
- Protocolo de investigación en RL
  - Diseño de algoritmos
  - Indicadores y evaluación
  - Comparación y comprobación de hipótesis
- Estado del arte de RL
- Resolver RL
- Realizar investigación RL
- Don't panic and don't forget your towel



# Objetivos del curso

## General

- Diseñar agentes capaces de mejorar con la experiencia en un problema dado

## Específicos

- Explicar los fundamentos de métodos de aprendizaje por refuerzo, sus alcances y limitaciones para solucionar problemas prácticos.
- Desarrollar y comparar métodos clásicos de planeación y control
- Resolver problemas donde no existe un modelo dado

# Problemas abiertos en RL

- En el núcleo de los algoritmos
  - Estabilidad: ¿converge?
  - Eficiencia: ¿cuánto toma para que converja?
  - Generalización: después de convergencia, ¿generaliza?
- En las suposiciones
  - ¿Es la formulación correcta del problema?
  - ¿Cuál es la fuente de supervisión?

# Estabilidad e hiper parámetros

- Desarrollar algoritmos de RL estables es complicado
- Q-learning
  - No garantizan convergencia
  - Muchos parámetros para estabilidad
- Política de gradiente
  - Alta varianza
  - Muchas muestras
  - Muchos parámetros
- Algoritmos basados en modelo
  - Métodos de ajuste
  - Optimización de la política no es trivial
  - Las políticas tienden a explotar el modelo



# Áreas de investigación activas

Fundamentos  
matemática

Multi-tarea

Meta-  
aprendizaje

Multi-agente

RL Bayesiano

Ambientes  
altamente  
complejos

Transferencia  
de  
aprendizaje

RL Multi-  
objetivo



# Guía general para publicar en ML

Novedad del algoritmo

Novedad de la aplicación/problema

Dificultad de la aplicación

Calidad de los resultados

Lecciones aprendidas

## Estilos de artículos para RL

Proponer un nuevo modelo de aprendizaje y algoritmo para ese modelo

Proponer un algoritmo que mejora al estado del arte en alguna forma

Probar que una tarea es difícil o imposible

Teoría

Aplicaciones

# Investigación en RL

## Teórica

- Discutir todas las suposiciones del trabajo
- Incluir demostraciones de los resultados teóricos
- Nuevos algoritmos con garantía matemática
- Demostraciones a cerca de algoritmos o formulaciones existentes

## Empírica

- Incluir código, datos e instrucciones para reproducir los resultados experimentales
- Datos de entrenamiento, hiper parámetros , etc.
- Reportar barras de error, intervalos de confianza, investigar para significancia estadística
- Nuevos algoritmos
- Comparación con algoritmos del estado del arte

## Aplicada

- Aplicaciones que no se podían hacer o se están escalando
- Técnicas únicas para aplicaciones populares
- Lecciones aprendidas desde la perspectiva de ML para diferentes aplicaciones
- Estudios analíticos que comparan diferentes métodos en varios metodos



# ¿Cómo presentar los resultados?



Pregunta de investigación



Planeación experimental



La tarea



Setup



Resultados/Visualizaciones

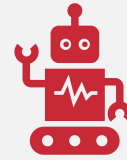


Observaciones

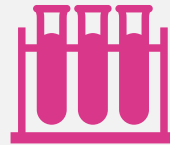


Discusión

# Garantizar reproducibilidad



Repetibilidad: mismo equipo, mismo setup experimental



Reproducibilidad: diferente equipo, mismo setup experimental



Replicabilidad: diferente equipo, diferente setup experimental

# ¿Cómo analizar los resultados?

## Exploración de datos

- Estadística descriptiva
- Visualizaciones: final, tiempo de ejecución

## Análisis de confirmación

- Pruebas estadísticas
  - Kruskal Wallis
  - Wilcoxon rank sum

## Análisis de relevancia

- Relevancia de la prueba utilizada

# Rankings para publicaciones

- <https://bidi.unam.mx/index.php/cobertura-tematica/ver-todos-los-recursos/555-journal-citation-reports-science-edition-full>
- <http://www.conferenceranks.com/>



Universidad Nacional  
Autónoma de México



[Inicio](#) [Colecciones digitales](#) [Herramientas de búsqueda](#) [Cobertura temática](#) [Redes Sociales Científicas](#) [Ayuda](#) [Iniciar sesión](#)

[Biblioteca Digital](#) > [Cobertura temática](#) > [Ver todas las Bases de Datos](#) > [Journal Citation Reports: Science Edition](#)

## Ver todas las Bases de Datos

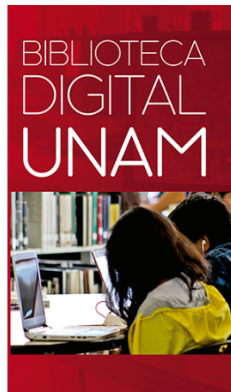


WEB OF SCIENCE™

Journal Citation Reports: Science Edition es una base de datos referencial que ofrece un medio sistemático y objetivo para evaluar de manera crítica las revistas más importantes del mundo en las áreas de ciencia y tecnología. Este recurso de evaluación de revistas, brinda información estadística basada en los datos de citas, al recopilar las referencias citadas (que suministran los propios autores de los artículos) y permite medir la influencia y el impacto de las investigaciones realizadas (a nivel de revistas y categorías). Asimismo, muestra las relaciones entre las revistas que citan y las que son citadas. La base de datos incluye más de 5,907 títulos de revistas.

Buscar en el sitio

Ej. acceso, vigencia



COMPUTER SCIENCE, ARTIFICIAL INTEL... SCIE JCR Year: 2020								
Journal name	ISSN	eISSN	Category	Total Citations	2020 JIF	JIF Quartile	2020 JCI	% of OA Gold
<input type="checkbox"/> IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	0162-8828	1939-3539	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	59,682	16.389	Q1	5.24	9.38 %
<input type="checkbox"/> Nature Machine Intelligence	N/A	2522-5839	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	1,212	15.508	Q1	3.14	0.00 %
<input type="checkbox"/> Information Fusion	1566-2535	1872-6305	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	9,059	12.975	Q1	4.34	3.48 %
<input type="checkbox"/> IEEE TRANSACTIONS ON FUZZY SYSTEMS	1063-6706	1941-0034	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	20,515	12.029	Q1	2.96	0.38 %
<input type="checkbox"/> IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION	1089-778X	1941-0026	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	16,918	11.554	Q1	3.09	2.25 %
<input type="checkbox"/> IEEE Transactions on Cybernetics	2168-2267	2168-2275	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	24,753	11.448	Q1	2.63	2.52 %
<input type="checkbox"/> IEEE Computational Intelligence Magazine	1556-603X	1556-6048	COMPUTER SCIENCE, ARTIFICIAL INTELLIGENCE - SCIE	2,039	11.356	Q1	1.45	3.03 %



# Grupos de investigación en México

IIMAS-  
UNAM

- Dr. Antonio Neme
- Dr. Carlos Hernández

INAOE

- Eduardo F. Morales

Cinvestav-  
IPN

- Dr. Ricardo Landa

IPN

- Ricardo Menchaca

INEEL

- Alberto Reyes Ballesteros

# Investigación en el extranjero

Reasoning and Learning Lab,  
McGill University,  
Canada (Doina Precup and  
Joelle Pineau are with  
DeepMind and FAIR Montreal  
respectively)

Reinforcement Learning and  
Artificial Intelligence, U Alberta,  
Canada (Rich Sutton, Michael  
Bowling, Patrick Pilarski are with  
DeepMind Edmonton; Csaba  
Szepesvári is with DeepMind  
London)

Reinforcement learning and  
online learning group, Imperial  
College London, UK (Marc  
Deisenroth is at Prowler.IO)

Whiteson Research Lab, U  
Oxford, UK

Inria SequeL, Lille,  
France (Mohammad  
Ghavamzadeh, Rémi Munos,  
Bilal Piot are at DeepMind,  
Alessandro Lazaric is at FAIR  
Paris, Olivier Pietquin is at  
Google Brain)

Juergen Schmidhuber's group,  
IDSIA, Switzerland (now at  
NNAISENSE)

Shie Mannor's group at  
Technion, Israel

Gergely Neu at UPF Barcelona,  
Spain

Balaraman Ravindran's group at  
IIT-Madras, India

Jimmy Ba's group at U Toronto/  
Vector Institute, Canada

Alireza Makhzani, Amir-  
Massoud Farahmand at Vector  
Institute, Canada

Deep Mind

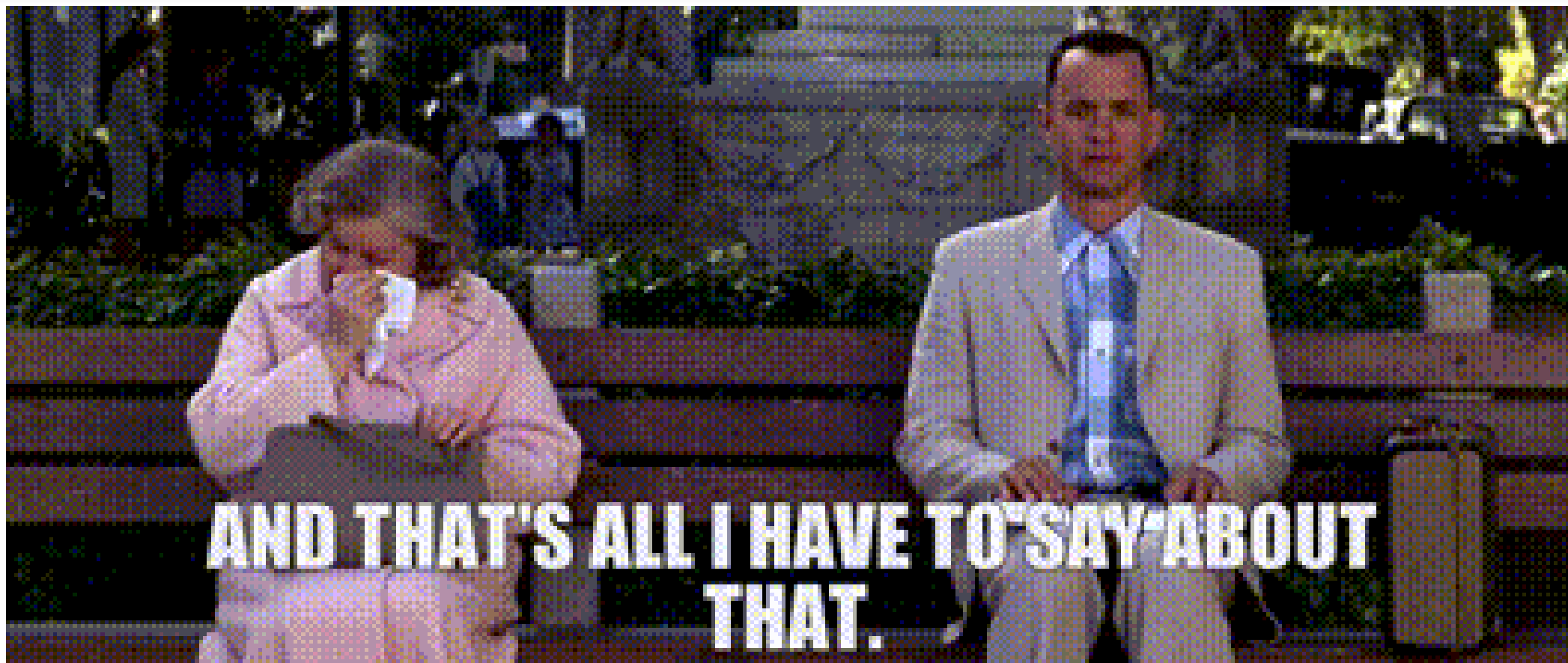
Robotic Artificial Intelligence  
and Learning Lab, Berkeley

Stanford Machine Learning  
Group, Stanford

# Código

- <https://github.com/openai/baselines>
- <https://github.com/DLR-RM/stable-baselines3>
- <https://github.com/dennybritz/reinforcement-learning>
- [https://github.com/omerbsezer/Reinforcement\\_learning\\_tutorial\\_with\\_demo#GridWorld](https://github.com/omerbsezer/Reinforcement_learning_tutorial_with_demo#GridWorld)

Y entonces...







¿Comentarios,  
dudas, quejas  
y sugerencias?

---







# Para la otra vez...

- Examen
- Proyecto
- Exámenes finales



The End.

