

08MIAR - Aprendizaje por refuerzo

Sesión 1 – Introducción

The logo consists of a solid orange circle containing the lowercase letters "viu" in a white, sans-serif font.

viu

Universidad
Internacional
de Valencia

De:

 Planeta Formación y Universidades

Sobre mí

Gabriel Muñoz

Lead Data Scientist en Intelygenz (<https://intelygenz.com/>)

Co-fundador de *MAD_RL*, comunidad de aprendizaje por refuerzo con base en Madrid

*gabrielenrique.munoz@campusviu.es
<https://es.linkedin.com/in/gabrielmunozrios>*

Sobre la asignatura

El contenido de la asignatura se organizará en tres bloques:

- En el primer bloque introduciremos el contexto en el que se desarrollan las soluciones de aprendizaje por refuerzo, así como los conceptos y términos más importantes.
- El segundo bloque estará compuesto del estudio de los algoritmos que forman el estado del arte actual. Principalmente, cuáles son las estrategias de aprendizaje y cómo se produce el *sampling* de datos durante el mismo.
- El tercer bloque cubrirá una serie de sesiones prácticas donde implementaremos soluciones de aprendizaje por refuerzo usando diferentes *frameworks* y librerías.

Para más información sobre las sesiones: **Anexo de organización de sesiones en la sección de guía didáctica de la asignatura.**

Sobre la asignatura

Respecto a los requisitos de la asignatura:

- Los conceptos teóricos que veremos se basarán en conceptos matemáticos y estadísticos. Hablaremos de distribuciones de probabilidad, de búsqueda óptima, de cadenas de *Markov*, etc. El conocer estos conceptos no es una restricción para el curso, pero si que es recomendable.
- En la parte práctica trabajaremos con Python y algunas librerías típicas de entornos para trabajar con datos, como por ejemplo *Keras*, *Tensorflow* y *Pytorch*. Habrá disponible una sesión práctica grabada enfocada en la puesta a punto del entorno y la presentación de los frameworks que utilizaremos.
- En ambas partes será imprescindible conocer y tener cierta experiencia con modelos *Deep Learning* y, en concreto, con redes convolucionales.

Sobre la asignatura

El método de evaluación de la asignatura estará compuesto de:

- Participación en foros de debate (10%)
- Actividad sobre artículo científico (10%)
- Actividad sobre una serie de vídeos teóricos (10%)
- Proyecto práctico (30%)
- Examen final (40%)

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

¿Qué es aprendizaje por refuerzo?

“Aprender a partir de la interacción es una idea fundamental y común a todas las teorías del aprendizaje y de la inteligencia”

“Aprendizaje por refuerzo [...] es a la vez un problema, un conjunto de soluciones que funcionan bien sobre un conjunto de problemas y el campo que estudia esos problemas y los métodos que pueden solucionarlo.”

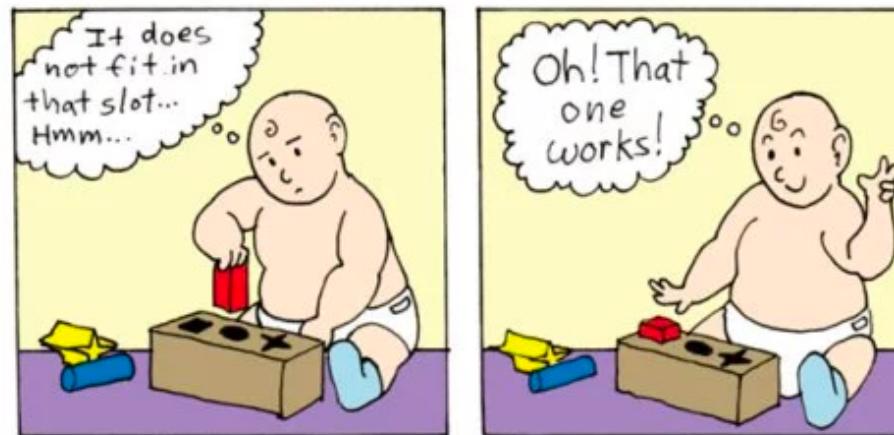
Reinforcement learning: An Introduction, R. Sutton & A. Barto

“El objetivo del aprendizaje por refuerzo es entrenar un agente inteligente que es capaz de interactuar con un entorno de manera inteligente.”

Deep Q Network vs Policy gradients, Felix Yu

¿Qué es aprendizaje por refuerzo?

Podemos ver el aprendizaje por refuerzo como el estudio y diseño de **agentes** que aprenden por medio de **prueba-y-error**.



<https://rochemamabolo.files.wordpress.com/2018/07/trial02.jpg?w=490>

Sesión 1 - Introducción

¿Qué es aprendizaje por refuerzo?

Si nos paramos y analizamos otras ramas dentro del aprendizaje basado en datos, encontramos dos grandes conjuntos: métodos supervisados y métodos no supervisados. Podríamos añadir aprendizaje por refuerzo como un tercer conjunto dentro de esta clasificación.

Desde un punto de vista de datos, podemos definir la relación de cada conjunto con el conocimiento que se puede obtener a partir de los datos:

Análisis descriptivo → Métodos no supervisados

Análisis predictivo → Métodos supervisados

Análisis prescriptivo → Métodos aprendizaje por refuerzo

¿Qué es aprendizaje por refuerzo?

Si nos paramos y analizamos otras ramas dentro del aprendizaje basado en datos encontramos dos grandes conjuntos: métodos supervisados y métodos no supervisados. Podríamos añadir aprendizaje por refuerzo como un tercer conjunto dentro de esta clasificación.

Desde un punto de vista de datos, podemos definir la relación de cada conjunto con el conocimiento que se puede obtener a partir de los datos:

Análisis descriptivo → Métodos no supervisados

Análisis predictivo → Métodos supervisados

Análisis prescriptivo → Métodos aprendizaje por refuerzo

¿Qué es aprendizaje por refuerzo?

Si nos paramos y analizamos otras ramas dentro del aprendizaje basado en datos encontramos dos grandes conjuntos: métodos supervisados y métodos no supervisados. Podríamos añadir aprendizaje por refuerzo como un tercer conjunto dentro de esta clasificación.

Desde un punto de vista de datos, podemos definir la relación de cada conjunto con el conocimiento que se puede obtener a partir de los datos:

Análisis descriptivo → Métodos no supervisados

Análisis predictivo → Métodos supervisados

Análisis prescriptivo → Métodos aprendizaje por refuerzo

¿Qué es aprendizaje por refuerzo?

Si nos paramos y analizamos otras ramas dentro del aprendizaje basado en datos encontramos dos grandes conjuntos: métodos supervisados y métodos no supervisados. Podríamos añadir aprendizaje por refuerzo como un tercer conjunto dentro de esta clasificación.

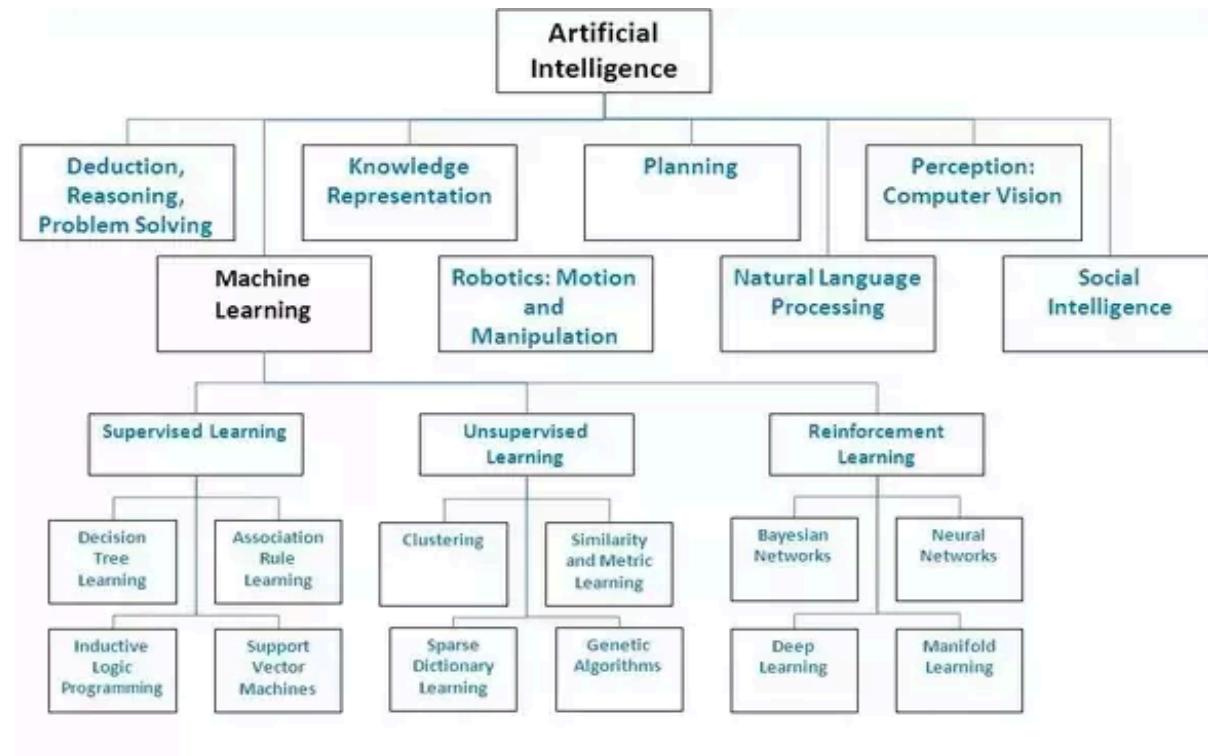
Desde un punto de vista de datos, podemos definir la relación de cada conjunto con el conocimiento que se puede obtener a partir de los datos:

Análisis descriptivo → Métodos no supervisados

Análisis predictivo → Métodos supervisados

Análisis prescriptivo → Métodos aprendizaje por refuerzo

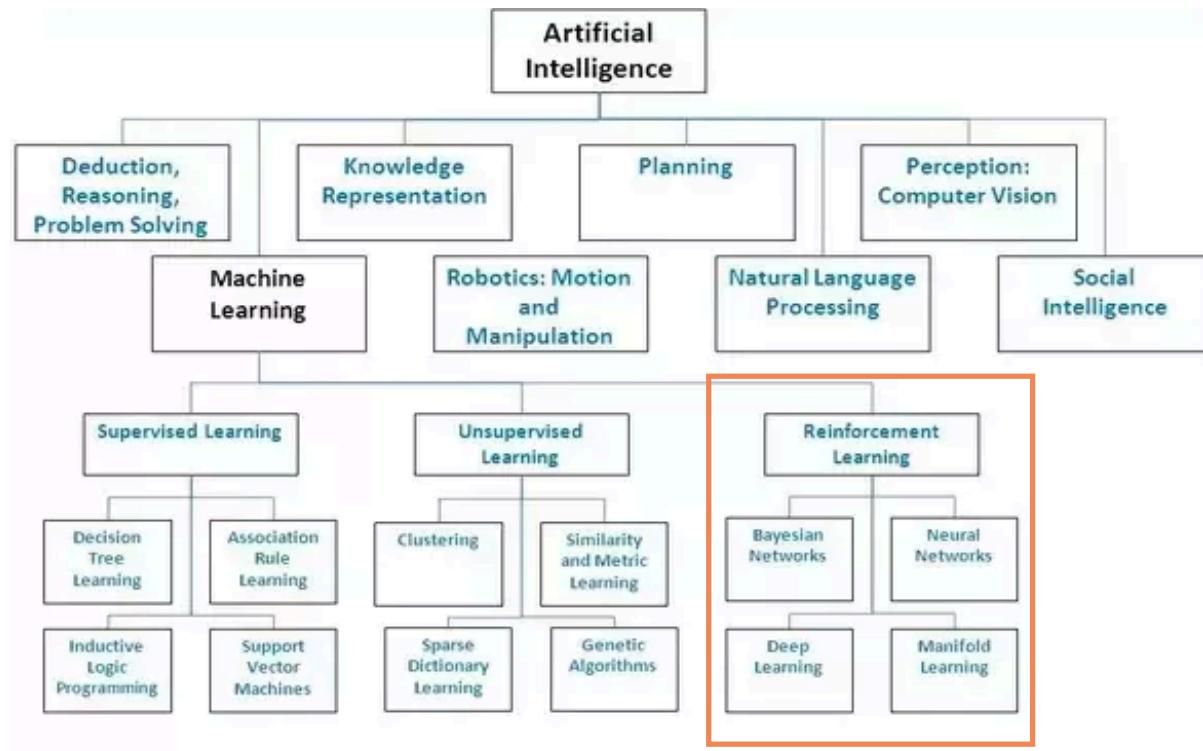
¿Qué es aprendizaje por refuerzo?



<https://qph.fs.quoracdn.net/main-qimg-541146e2d9d611a42ebb5074aa72fef4.webp>

Sesión 1 - Introducción

¿Qué es aprendizaje por refuerzo?



<https://qph.fs.quoracdn.net/main-qimg-541146e2d9d611a42ebb5074aa72fef4.webp>

Sesión 1 - Introducción

¿Qué es aprendizaje por refuerzo?

En mi opinión, el aprendizaje por refuerzo desarrolla de una manera natural e intuitiva lo que conocemos (o interpretamos) como aprendizaje.

Al producirse el aprendizaje de una manera similar a la de los humanos, se podría empezar a hablar y discutir en términos de adaptación y/o transferencia de conocimiento entre distintos problemas.

Precisamente la transferencia de conocimiento entre distintos problemas (***transfer learning***) es una rama de investigación con mucha fuerza hoy en día dentro de la inteligencia artificial y, en concreto, del ***Deep Learning***.

Otro punto muy interesante es el concepto de bias o sesgo, que tanto impacto tiene en los datos con los que trabajamos. Dentro de una solución guiada por aprendizaje por refuerzo podríamos mitigar mucho este problema, ya que el aprendizaje se produce *from scratch*.

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Estado del arte

Como en otras ramas de la inteligencia artificial, los primeros algoritmos y soluciones basadas en aprendizaje por refuerzo datan de hace algunas décadas. En nuestro caso, fue en los años 80/ principios de los 90 cuando hubo una tendencia positiva en el estudio y uso de algoritmos basados en aprendizaje por refuerzo.

Aunque los resultados fueron prometedores, debido a las capacidades computacionales y al desarrollo de otros algoritmos con enfoques similares en aquel momento (**algoritmos evolutivos**), las soluciones basadas en aprendizaje por refuerzo disminuyeron hasta hace unos años.

Como apunte, la principal diferencia entre algoritmos de aprendizaje por refuerzo y algoritmos evolutivos es que los agentes de algoritmos evolutivos no aprenden de la interacción con su entorno. Las reglas de su interacción son inamovibles, por lo que no pueden aprovechar todo la información de la que disponen para aprender y adaptarse al problema que intentan resolver.

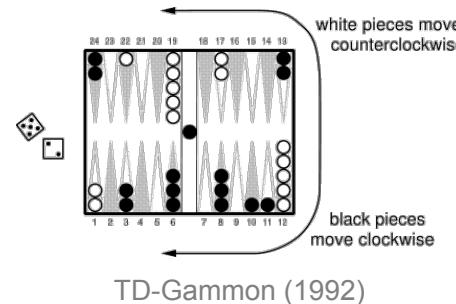
Estado del arte

Como decíamos, hasta hace algunos años no hubo una explosión en este tipo de algoritmos y del potencial que tienen. Una relación directa ha sido el magnífico desarrollo de las técnicas de *Deep Learning* así como la capacidad computacional que ahora tenemos a nuestra disposición.

Igualmente, si tuviéramos que elegir tres hitos que sentaron las bases para hablar de aprendizaje por refuerzo hoy serían:



Minsky PHD thesis (1954)



TD-Gammon (1992)



DeepMind (2015)

Estado del arte

Vamos a centrarnos en Google Deepmind. El gran hito que consiguió Google Deepmind con su solución fue que combinó *Deep Learning* con Aprendizaje por refuerzo.

Hasta ese momento, las variables que se usaban en algoritmos por refuerzo eran variables *tabuladas* o definidas manualmente. Con la solución de DeepMind, se usó una red neuronal como función aproximadora para que el agente tomara decisiones, algo inaudito hasta el momento y con una base de inteligencia artificial formidable.

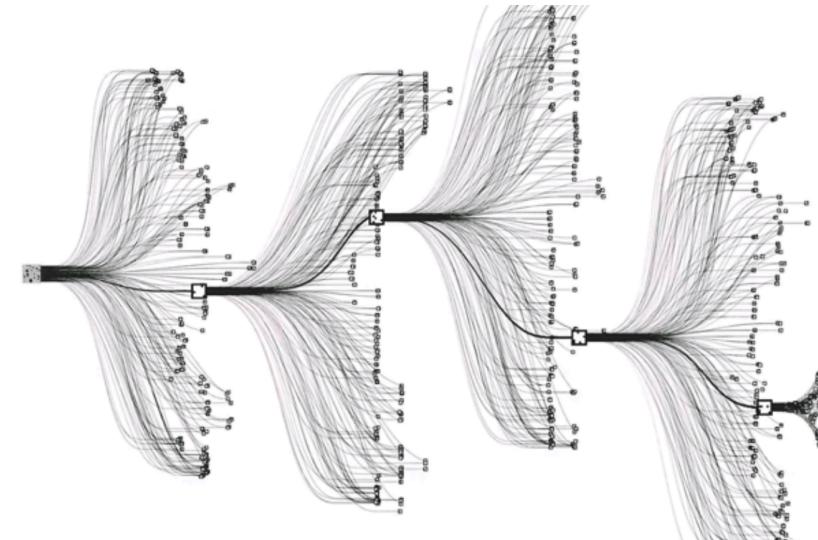
Este primer hito se realizó sobre un entorno de simulación de videojuegos de Atari.



https://deepmind.com/static/v0.0.0/images/deepmind_logo.png

Sesión 1 - Introducción

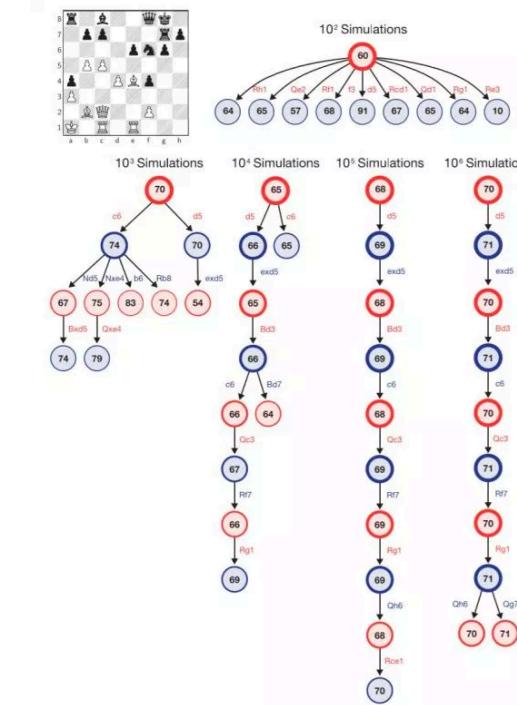
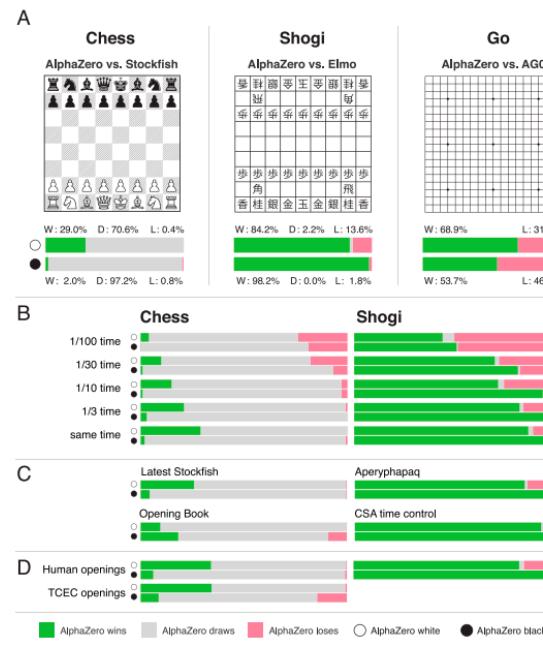
Estado del arte



Hitos de DeepMind: **AlphaGo**

https://obs-educom.cdnstatics.com/sites/default/files/styles/blog_post/public/post/blog_obs_tendencias_e_innovacion_roman-febrero_1_foto3.jpg?itok=caNOUFa-
https://cdn-images-1.medium.com/max/2600/1*qpzAxoUR9POLY1_zJhU5g.png

Estado del arte



Hitos de DeepMind: **AlphaZero**

<https://cdn.chess24.com/f1DREmrdrFOR-HKtB3JRcA/original/chess-performance.png>
<https://vonneumannmachine.files.wordpress.com/2018/12/captura.jpg?w=614&h=893>

Estado del arte

Otra compañía que ha sido un pulmón en el desarrollo del aprendizaje por refuerzo estos últimos años ha sido OpenAI.

La misión de OpenAI es la creación de una inteligencia artificial de carácter general con una fuerte base ética y de principios.

OpenAI han sido los responsables de algunos de los algoritmos que ahora mismo son estado del arte, así como de ser la primera compañía en ganar con inteligencia artificial en entornos de colaboración (y muy avanzados en cuanto a complejidad). Un ejemplo es el videojuego *DOTA2*.



<https://venturebeat.com/wp-content/uploads/2019/03/openai-1.png?fit=2400%2C1000&strip=all>

Sesión 1 - Introducción

Estado del arte

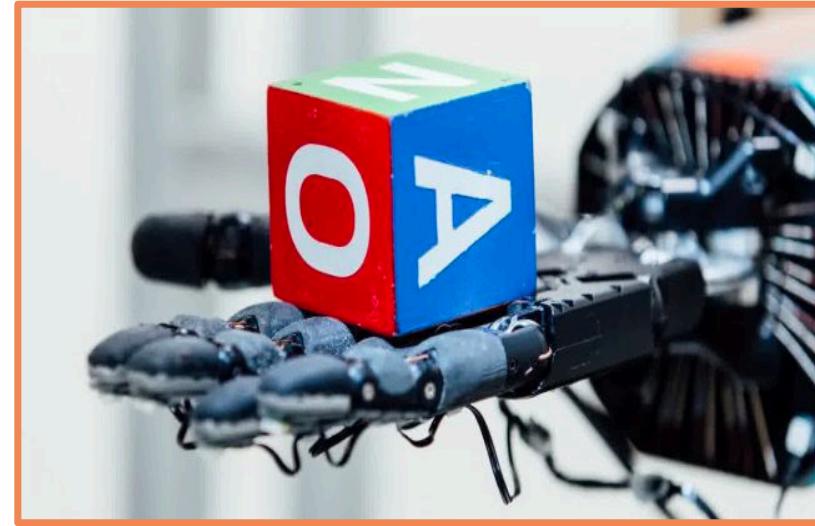
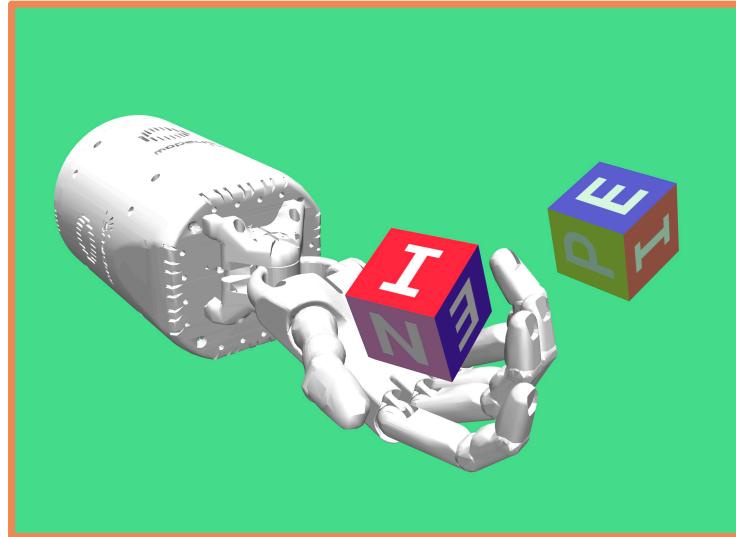


Hitos de OpenAI: Dota2

<https://cdn.estnn.com/wp-content/uploads/2019/04/22152826/Dota-2-Open-AI-2.png.jpg>

Sesión 1 - Introducción

Estado del arte



Hitos de OpenAI: **Robótica**

<https://spectrum.ieee.org/image/MzAyMzczMg.jpeg>
<https://www.extremetech.com/wp-content/uploads/2018/07/openai-640x353.jpg>

Sesión 1 - Introducción

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Distintos enfoques del aprendizaje por refuerzo

Una posible clasificación que podemos hacer cuando hablamos de aprendizaje por refuerzo está relacionado con el entorno donde se ejecutará la solución desarrollada.

En este sentido tenemos dos grupos principales:

Entorno puramente digital - Ámbito de investigación

Entorno Digital-a-Real - Ámbito de investigación/aplicación

Vamos a entrar en detalle con cada enfoque.

Distintos enfoques del aprendizaje por refuerzo



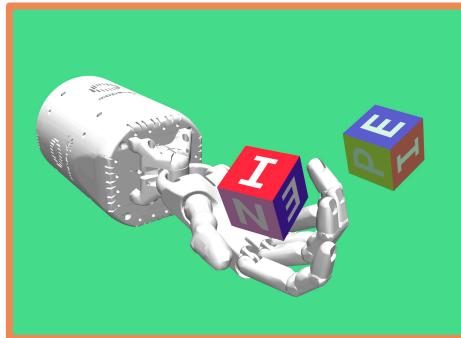
Entorno puramente digital - Ámbito de investigación

Este es el entorno típico que nos encontraremos en la mayoría de ejemplos a día de hoy.

Trabajaremos en un entorno totalmente simulado y controlado para poder probar nuevos algoritmos, nuevos modelos, etc.

Este enfoque es fundamental para contrastar hipótesis teóricas y comprobar su posible aplicación en el mundo real.

Distintos enfoques del aprendizaje por refuerzo



Entorno Digital-a-Real - Ámbito de investigación

En este caso tenemos una situación parecida al enfoque anterior aunque ahora usaremos el conocimiento extraído en un entorno real.



Un caso común es la robótica, en el que se diseña toda la simulación en un entorno digital para usar la solución obtenida en brazos robóticos reales.

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Retos actuales y de futuro

El principal reto en el que el aprendizaje por refuerzo se centra actualmente es encontrar **casos de uso** en los que se pueda aplicar de una manera beneficiosa.

Es verdad que en los últimos años se han ido desarrollando soluciones con aplicaciones reales en ámbitos como la robótica, control automático, etc., pero todavía no tienen un impacto de negocio como para tomar a estas soluciones más en “serio”.

Para la mayoría de estos casos hay soluciones tradicionales que funcionan muy bien y no necesitan de la complejidad que necesita el aprendizaje por refuerzo.

Retos actuales y de futuro

Una de las complejidades que presentan las soluciones basadas en aprendizaje por refuerzo, y que veremos en las siguientes sesiones, es **la cantidad de hiperparámetros** que se necesitan configurar para que una solución funcione.

Y no sólo eso, ¿qué significa que una *solución funcione*?

Veremos en siguientes sesiones conceptos para poder discutir sobre esta pregunta, pero lo que está claro es que el concepto de que algo sea útil o no sigue estando muy sesgado a los ojos de quien lo mire (como es normal por otra parte).

Retos actuales y de futuro

Otro reto es que aún teniendo mucho dominio de la materia y conociendo muy bien todos los conceptos, el control sobre el proceso de aprendizaje es vago.

En otras disciplinas de la inteligencia artificial, como por ejemplo en la aplicación de técnicas de Deep Learning, tenemos una metodología que, aunque sea empírica, podemos controlar y medir cómo el sistema evoluciona y si lo está haciendo de forma favorable o no.

En el caso del aprendizaje por refuerzo no se dan las mismas circunstancias. Dependiendo del algoritmo, del reto y de la información disponible, la forma de medir la bondad de una solución difiere y por tanto le añade una complejidad extra a su diseño.

Retos actuales y de futuro

Aún así, todos los grupos de investigación comparten la idea de que el **potencial que tiene este enfoque de la inteligencia artificial** es *infinito*.

La mayoría de problemas que ahora mismo podemos solucionar son problemas más cercanos a la automatización que al razonamiento y adaptación que una inteligencia real permitiría.

Es aquí donde también se presenta un gran reto: ser capaz de encontrar la madurez suficiente para afrontar verdaderos problemas de Inteligencia Artificial siguiendo esta metodología.

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

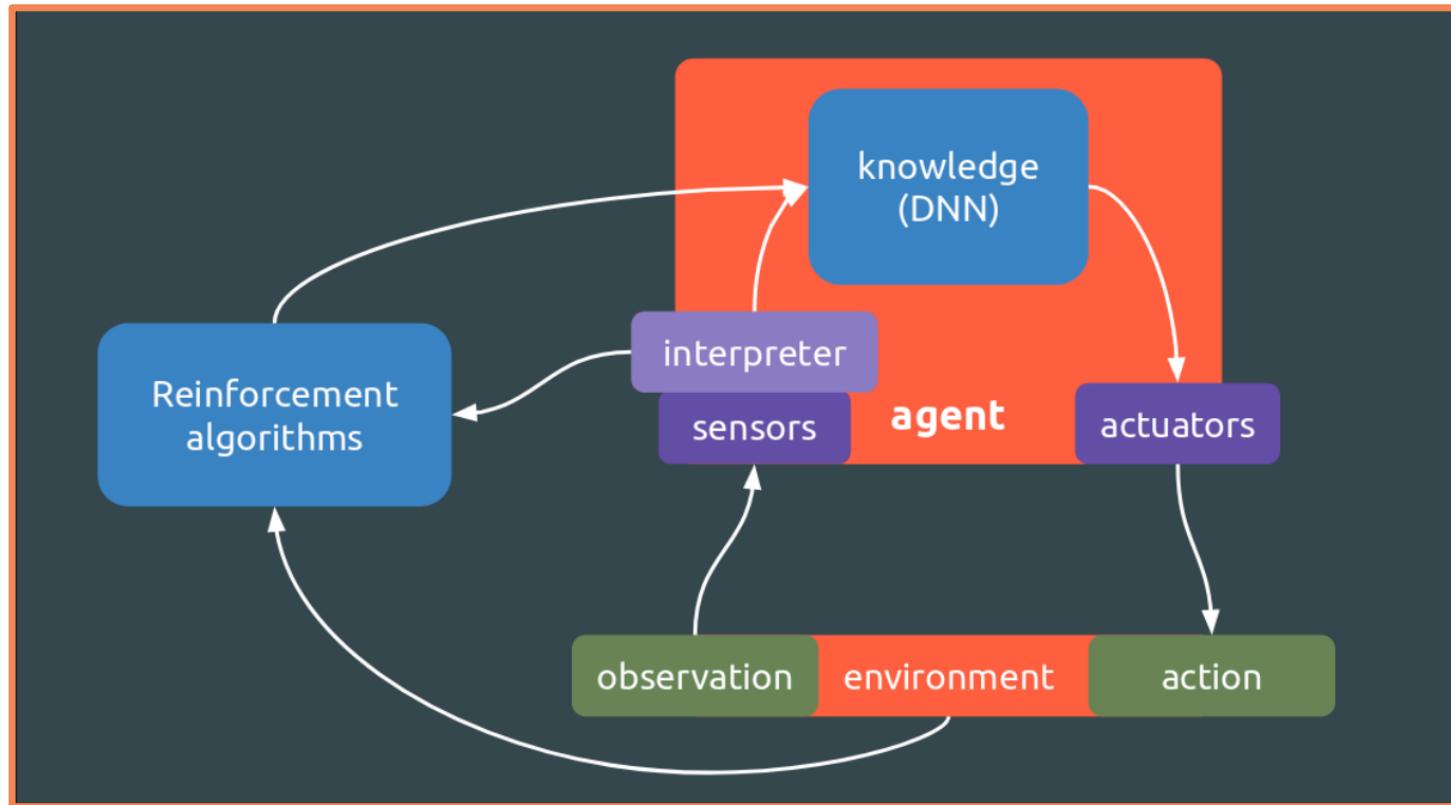
Retos actuales y de futuro

Vista general

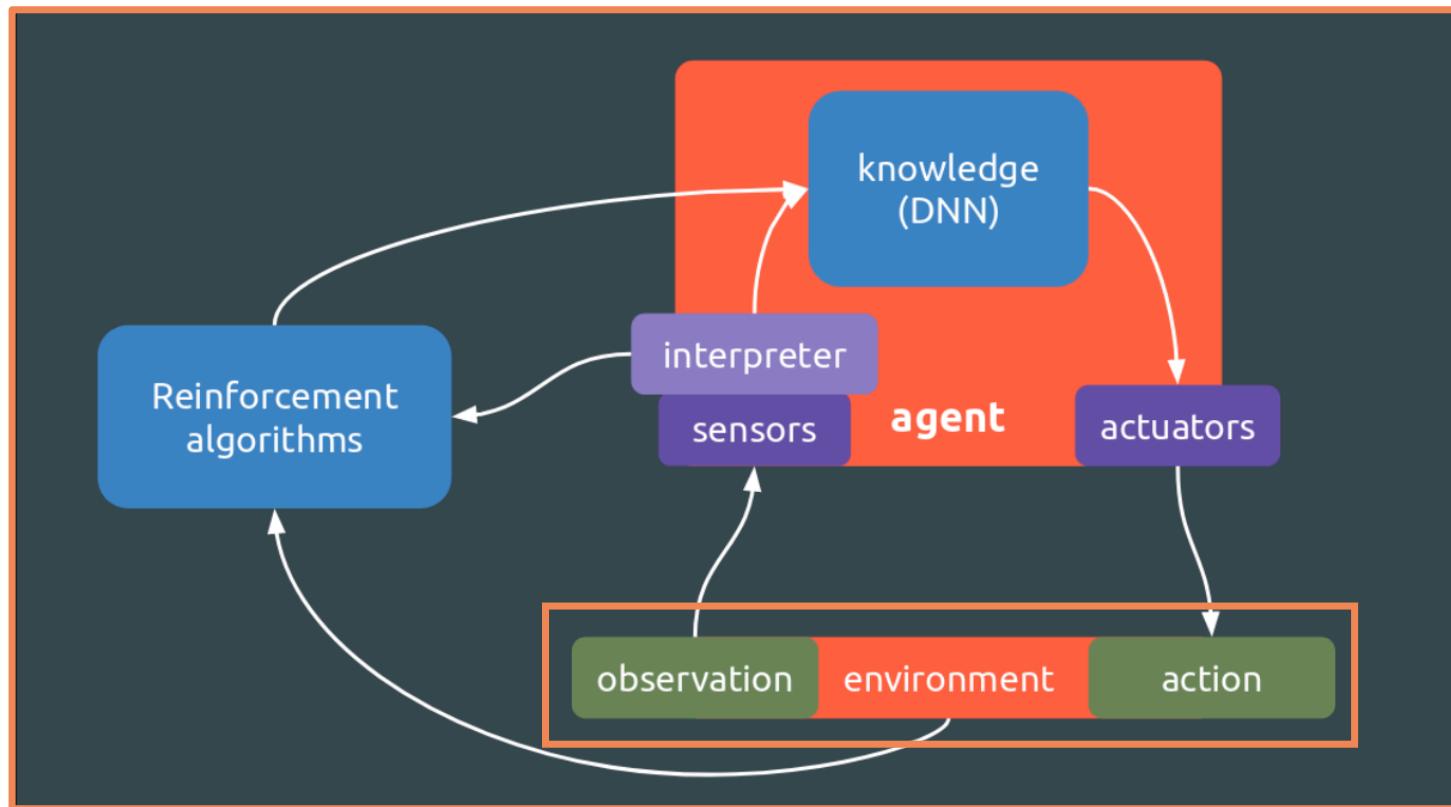
Conclusiones

Bibliografía recomendada

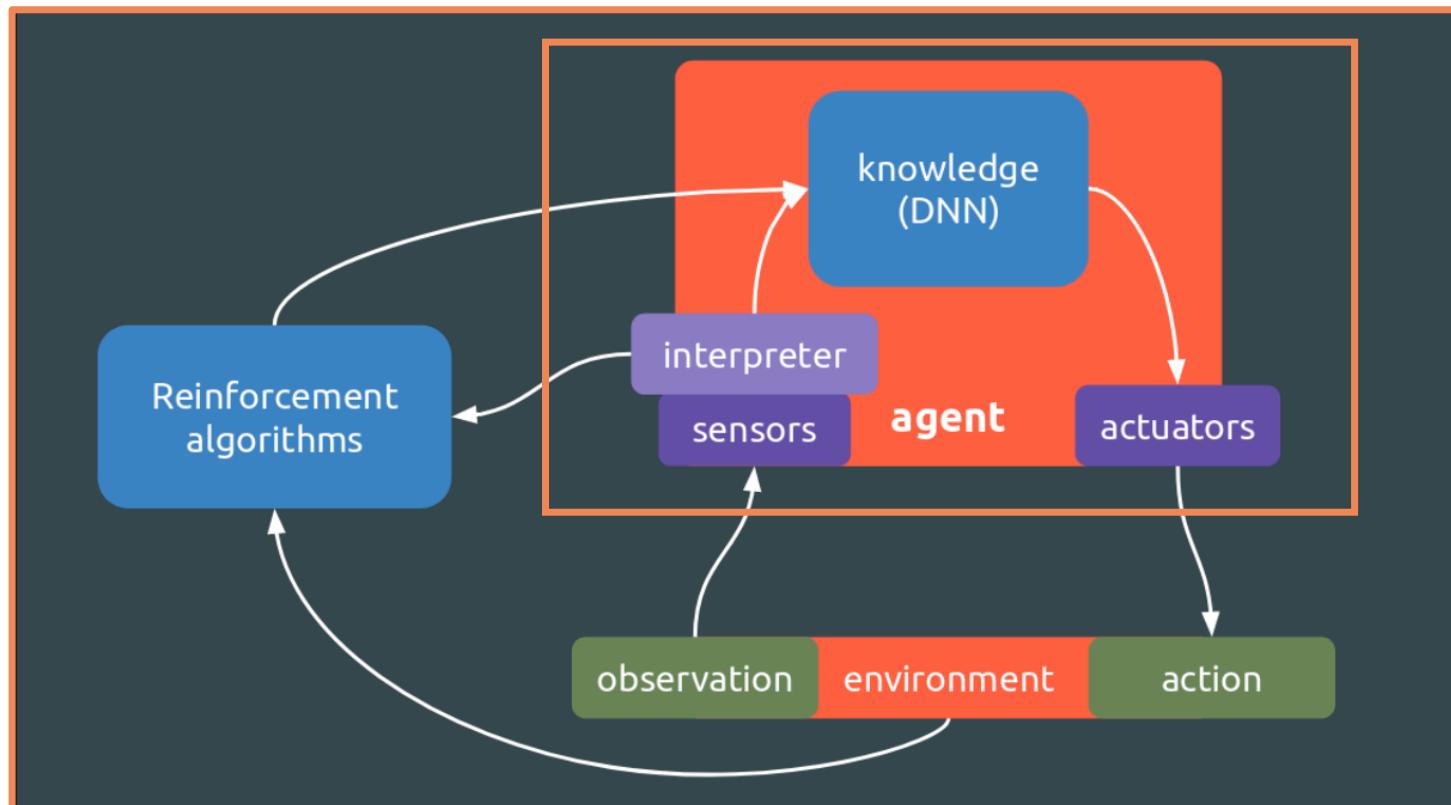
Vista general



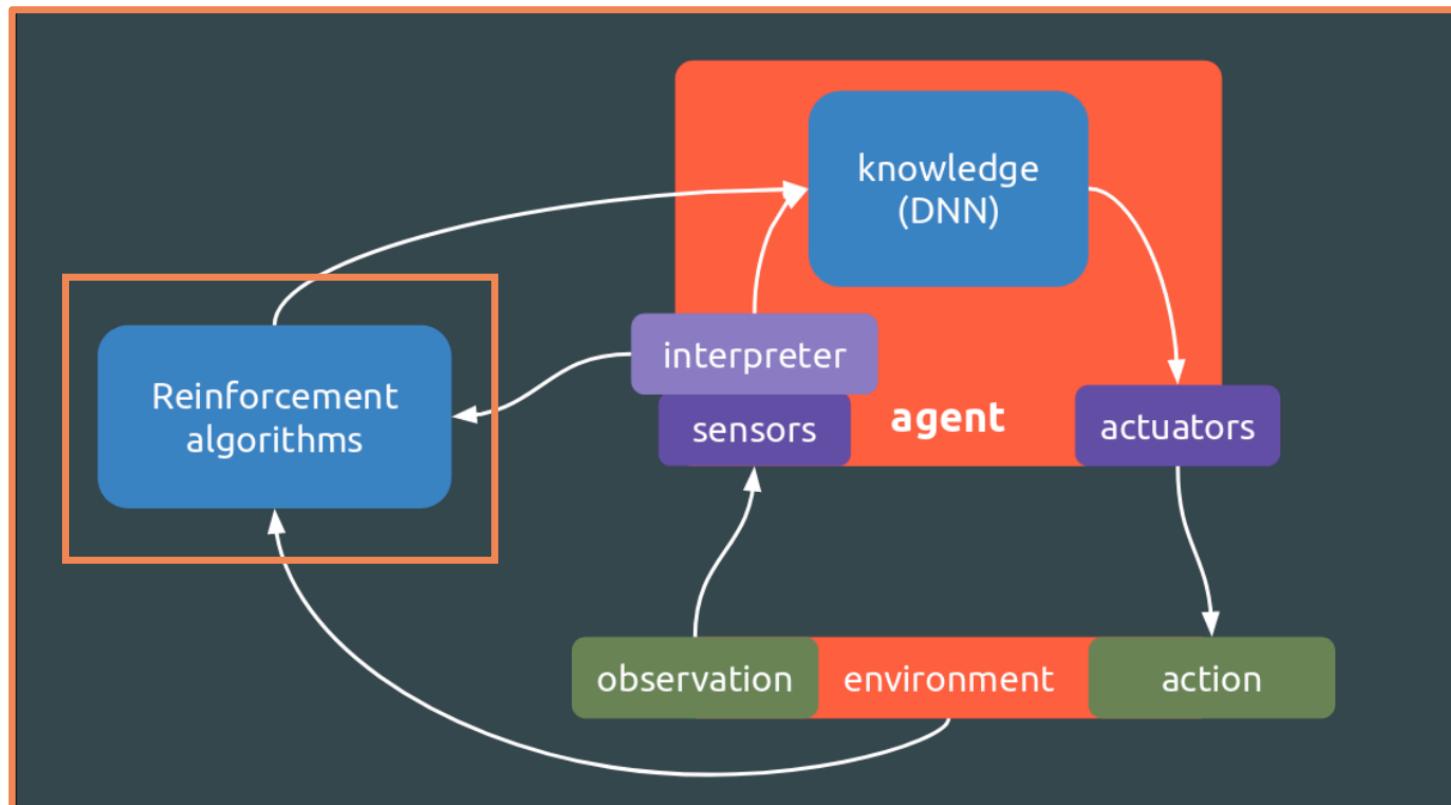
Vista general



Vista general



Vista general



Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Conclusiones

- El aprendizaje por refuerzo es una de las ramas más prometedoras dentro de la inteligencia artificial.
- El proceso de aprendizaje se asemeja a la idea de aprendizaje que tenemos los seres humanos.
- Aunque es una rama muy prometedora, todavía están en sus primeros pasos. No está totalmente asentada y sus soluciones no se pueden aplicar en todos los campos y/o dominios.
- La unión de modelos de Deep Learning con algoritmos de aprendizaje por refuerzo ha sido la mezcla perfecta para abrir todo el conjunto de posibilidades con tanto potencial que tenemos a nuestro alcance actualmente.

Índice

¿Qué es aprendizaje por refuerzo?

Estado del arte

Distintos enfoques de aprendizaje por refuerzo

Retos actuales y de futuro

Vista general

Conclusiones

Bibliografía recomendada

Bibliografía recomendada

“Reinforcement Learning: An introduction”, Sutton y Barto:
<http://incompleteideas.net/book/bookdraft2017nov5.pdf>

Contenidos y documentación de OpenAI SpinningupRL:
<https://spinningup.openai.com/en/latest/>



viu

Universidad
Internacional
de Valencia

universidadviu.com

De:
 Planeta Formación y Universidades