



11°
Congreso
TUCUMÁN
2023

20
23



CoNaII SI

XI Congreso Nacional
de Ingeniería en Informática
y Sistemas de Información



XI Congreso Nacional de Ingeniería en Informática y Sistemas de Información.

CoNaISI 2023

Universidad Tecnológica Nacional.
Facultad Regional Tucumán

XI Congreso Nacional de Ingeniería en Informática y Sistemas de Información :
CoNalSI 2023 / Gabriela Ivana Aparicio ... [et al.] ; Compilación de José Augusto
Nasrallah ; Gabriela Ivana Aparicio ; Editado por Fernando Cejas. - 1a ed - Ciudad
Autónoma de Buenos Aires : edUTecNe, 2024.
Libro digital, PDF

Archivo Digital: descarga y online
ISBN 978-987-8992-38-9

1. Ingeniería. 2. Sistemas de Información. 3. Tecnología Informática. I. Aparicio, Gabriela
Ivana, comp. II. Nasrallah, José Augusto, comp. III. Cejas, Fernando, ed.
CDD 629.801

Diseño de Tapa e interior: Fernando Cejas



Universidad Tecnológica Nacional – República Argentina
Rector: Ing. Rubén Soro
Vicerrector: Ing. Haroldo Avetta
Secretaría Cultura y Extensión Universitaria: Ing. Federico Olivo Aneiros



Universidad Tecnológica Nacional – Facultad Regional Tucumán
Decano: Mgter Ing Walter Fabián Soria
Vicedecano: Ing. Juan Esteban Campos
Secretario de Vinculación Tecnológica: Ing. Mario Augusto Madariaga
Secretario de relaciones institucionales: Ing. Jorge Molina



edUTecNe – Editorial de la Universidad Tecnológica Nacional
Coordinador General a cargo: Fernando Cejas
Dirección General: Mg. Claudio Véliz
Dirección de Cultura y Comunicación: Ing. Pablo Lassave

Queda hecho el depósito que marca la Ley N° 11.723

© edUTecNe, 2024

Sarmiento 440, Piso 6 (C1041AAJ)

Buenos Aires, República Argentina

Publicado Argentina – Published in Argentina

ISBN 978-987-8992-38-9



Reservados todos los derechos. No se permite la reproducción total o parcial de esta obra, ni su incorporación a un sistema informático, ni su transmisión en cualquier forma o por cualquier medio (electrónico, mecánico, fotocopia, grabación u otros) sin autorización previa y por escrito de los titulares del copyright. La infracción de dichos derechos puede constituir un delito contra la propiedad intelectual.

Transferencia de Conocimiento en Sistemas Multi-Agente Entrenados con Aprendizaje por Refuerzo: un Mapeo Sistemático de la Literatura

Mario Rivas

Universidad Abierta Interamericana
Facultad de Tecnología Informática
marioaugusto.rivascastillo@alumnos.uai.edu.ar

Carlos Neil

Universidad Abierta Interamericana
Facultad de Tecnología Informática
carlos.neil@uai.edu.ar

Resumen

Contexto: El resultado del proceso de aprendizaje, en el ambiente tradicional del aprendizaje por refuerzo de sistemas multi-agente, no es directamente replicable en otros sistemas distintos al utilizado para el propio aprendizaje. El problema que surge a partir de esto es que, excepto casos de alta similitud entre ambientes, los sistemas multi-agente precisan comenzar nuevamente el proceso de aprendizaje desde cero. Objetivos: La contribución del presente trabajo consiste en un mapeo sistemático de los trabajos recientes relativos al proceso, técnicas y experimentación de transferencia de aprendizaje, en sistemas multi-agente. Método: se define un objetivo central de estudio y a partir del mismo, las preguntas de investigación que orientan el trabajo. A partir de esto se generan las cadenas de búsqueda, las cuales son iteradas en varias ocasiones para generar el universo de trabajos inicial. Luego de aplicar criterios de inclusión y exclusión, se procede a analizar cada uno de los mismos y se aplica una valoración cualitativa para reducir el grupo objetivo, del cual se generan las tablas resumen comparativas y las conclusiones. Resultados: se enumeran las principales técnicas utilizadas para la transferencia de conocimiento en sistemas multi-agente, destacando aquellas con mayor caudal de aportes investigativos y experimentales, así como los principales factores de complejidad mencionados y los contextos de aplicación de las mismas, tanto virtuales como materiales.

1. Introducción

El objetivo de este trabajo es la realización de un mapeo sistemático de la literatura académica relativa a los métodos y técnicas de transferencia de conocimiento (*Transfer Learning*), en el ámbito de los sistemas multi-agente (*Multi-Agent Systems*) con aprendizaje por refuerzo (*Reinforcement Learning*). Se define como “Sistema multi-agente” al dominio de investigación que se ocupa de sistemas compuestos por entidades de nivel micro (agentes) que poseen comportamiento autónomo y proactivo e interactúan con el ambiente, produciendo un comportamiento sistémico observable a nivel macro [1].

Son ampliamente utilizados en contextos físicos (tales como fábricas robotizadas, autos autónomos, casas inteligentes, etc.) como en contextos virtuales (sistemas de logística, gestión del tráfico, etc.), siendo su modelo de aprendizaje basado en reglas (modelo estático) o en técnicas de aprendizaje de máquina (*Machine Learning*). Dentro de ésta última, cobran especial relevancia las técnicas de aprendizaje por refuerzo, debido a la posibilidad de aprender empíricamente, mediante la experimentación y retro-alimentación.

El objetivo del agente con aprendizaje por refuerzo es maximizar la recompensa, por medio de la elección de las acciones que ofrezcan la mayor probabilidad de éxito en el estado actual. El concepto intuitivo de “recompensa”, inherente al ambiente, es fundamental para el aprendizaje por refuerzo. Valores mayores de esta función escalar, dependiente del estado actual y de la acción ejecutada, indican una conducta eficiente en pro de alcanzar los objetivos, en cuanto valores menores representan comportamientos no aconsejables [2].

Uno de los principales desafíos de este modelo de aprendizaje es la definición del balance entre “explorar” (el ambiente) y “explotar” (el conocimiento adquirido). Ante cada ciclo de experimentación, el agente precisa determinar si va a tomar una acción exploratoria (de la cual se carece de información previa) o si va a tomar una decisión basada en el conocimiento anteriormente formado. El agente puede “explotar” lo que ya conoce para obtener la recompensa conocida pero también precisa “explorar” para poder tomar mejores decisiones en el futuro [2]. A los efectos de encontrar la política de aprendizaje adecuada y obtener una base de conocimiento satisfactoria, se requiere un volumen de datos elevado, el cual difícilmente se puede obtener mediante experimentación en ambientes físicos, como por ejemplo el necesario para entrenar un auto con conducción autónoma. En dichos casos la solución para por la utilización de simuladores virtuales que, dependiendo de la complejidad del ambiente, pueden representar la mayor parte de la casuística real.

Independientemente del modelo de entrenamiento adoptado, resulta evidente la necesidad de compartir el aprendizaje entre agentes homogéneos (aquellos que poseen estructura sensorial y capacidad de acción idénticas) con el objetivo de replicar el conocimiento y reducir el tiempo de entrenamiento. Este proceso,

genéricamente definido como “Transferencia de Aprendizaje” o “Transferencia de Conocimiento”, ha sido objeto de decenas de trabajos académicos en los últimos años, orientados a proponer alternativas que resuelvan los desafíos principales. De acuerdo a [3] un problema remanente es la generalización de los métodos con el objetivo de permitir un rápido y eficiente aprendizaje multitarea. La idea de usar el conocimiento acumulado para este tipo de problemas es tomada de las habilidades del aprendizaje humano, que trabaja de forma muy similar. La transferencia de aprendizaje surgió de la necesidad de resolver este problema, y muchos investigadores se han focalizado en expandir los algoritmos de aprendizaje de máquina para que tengan la capacidad de transferir conocimiento, experiencia o habilidades, de forma que el aprendizaje subsiguiente requiera muy pocos ejemplos.

1. Trabajos Relacionados

Si bien no fue posible constatar la existencia de revisiones sistemáticas de la literatura que cubriesen los ejes temáticos propuestos en su totalidad, es importante destacar el estudio [4] y su contribución en este trabajo. La revisión sistemática propuesta por los autores aporta una visión estructurada de los métodos de transferencia de conocimiento disponibles a la fecha de su publicación (2019), que puede ser la base para definir una ontología para esta área. La descripción y el análisis de las técnicas conocidas así como de las herramientas para experimentación relacionadas, fueron utilizados como soporte de consulta para el presente estudio.

2. Preguntas de Investigación

Considerando el objetivo central de este trabajo y el marco temático, fueron definidas las siguientes preguntas de investigación:

Tabla 1. Preguntas guía del mapeo

PREGUNTAS DE INVESTIGACIÓN	MOTIVACIÓN
P1. ¿Cuáles son los principales métodos de transferencia de conocimiento?	M1. Detectar las principales líneas de investigación respecto a la transferencia de conocimiento.
P2. ¿Cuáles son los factores que impactan en la complejidad del proceso de transferencia del conocimiento?	M2. Identificar la casuística más relevante en la resolución del problema central.
P3. Desde el punto de vista de la aplicabilidad, ¿en qué contextos se plantea el trabajo, tanto físicos como	M3. Determinar la universalidad de las conclusiones arribadas por los investigadores y su

virtuales?	marco teórico-práctico.
P4. ¿Cuáles son las principales técnicas y herramientas utilizadas?	M4. Describir el conjunto de las técnicas y herramientas más frecuentemente utilizadas en este segmento temático.

3. Método de Revisión

El método utilizado en el presente trabajo consta de los siguientes pasos: (1) selección de repositorios de búsqueda, (2) definición de las palabras claves y generación de las cadenas de búsqueda, (3) proceso iterativo de ejecución de las búsquedas y ajuste de la cadena, (4) aplicación de criterios de inclusión/exclusión, (5) lectura y selección de estudios primarios y (6) preparación y síntesis de los datos extraídos.

4.1 Selección de Repositorios

Considerando la cantidad de documentos disponibles y la idoneidad con los temas estudiados, las fuentes seleccionadas fueron:

Tabla 2. Repositorios utilizados

IEEE Xplore Digital Library
ACM Digital Library
ResearchGate
Semantic Scholar

4.2 Palabras Clave y Términos Alternativos

Una vez determinados los repositorios de búsqueda, se listaron las palabras clave y algunos términos alternativos:

Tabla 3. Términos utilizados para la búsqueda en inglés

TÉRMINOS PRINCIPALES	TÉRMINOS ALTERNATIVOS
Reinforcement Learning	Artificial Neural Networks
	Machine Learning
	Reward Based Learning
Multi-Agent Systems	MAS
	MARL
Transfer Learning	Knowledge Transfer
	Generalization
	Distillation

Tabla 4. Términos utilizados para la búsqueda en castellano

TÉRMINOS PRINCIPALES	TÉRMINOS ALTERNATIVOS
----------------------	-----------------------

Aprendizaje por Refuerzo	Redes Neuronales Artificiales
	Aprendizaje de Máquina
	Aprendizaje Basado en Recompensas
Sistemas Multi-Agente	MAS
	MARL
Transferencia de Aprendizaje	Transferencia de Conocimiento
	Generalización
	Destilación

Tabla 5. Términos utilizados para la búsqueda en portugués

TÉRMINOS PRINCIPALES	TÉRMINOS ALTERNATIVOS
Aprendizado por Reforço	Redes Neurais Artificias
	Aprendizado de Máquina
	Aprendizado baseado em recompensa
Sistemas Multi-Agente	MAS
	MARL
Transferência de Aprendizado	Transferência de Conhecimento
	Generalização
	Destilação

4.3 Cadena de Búsqueda

En base a estos términos, y luego de un proceso iterativo de refinamiento de la cadena, se llegó a la cadena final de búsqueda, que incluyó también una versión en castellano y otra en portugués.

4.3.1 Cadena de Búsqueda en inglés

("Knowledge Transfer") AND ("Multi-Agent Systems") AND ("Reinforcement Learning")

4.3.2 Cadena de Búsqueda en castellano

("Transferencia de Conocimiento") AND ("Sistemas Multi-agente") AND ("Aprendizaje por Refuerzo")

4.3.3 Cadena de Búsqueda en portugués

("Transferência de Conhecimento") AND ("Sistemas Multi-agente") AND ("Aprendizado por Reforço")

4.4 Criterios de Inclusión y Exclusión

El total de trabajos inicialmente impactados por esta cadena en las fuentes fue tres mil cuatrocientos y ocho (3408), entre los cuales existían diversidad de tipos (artículos, tesis, libros, etc.) y amplitud de períodos (desde

1998 en adelante). A los efectos de reducir el listado de trabajos impactados y focalizar en el objetivo central, se aplicaron los siguientes criterios de inclusión y exclusión:

Tabla 6. Criterios de Inclusión y Exclusión

CRITERIOS DE INCLUSIÓN
CI 1. Estudios relacionados con transferencia de conocimiento en el ámbito del aprendizaje por refuerzo.
CI 2. Campo de investigación: ingeniería informática.
CI 3. Trabajos en inglés, español o portugués.
CI 4. Tipo de documento: documento de conferencia o artículo.
CRITERIOS DE EXCLUSIÓN
CE 1. Trabajos anteriores a 2010
CE 2. Se descartan las publicaciones que no usen algoritmos para el aprendizaje por refuerzo
CE 3. Trabajos de carácter preliminar

4.5 Lectura y Selección de Textos Primarios

En base a la aplicación de estos criterios, la cantidad de artículos seleccionados se redujo a trescientos cuarenta y tres (343). El siguiente paso fue el análisis de los resúmenes (*abstracts*) con el objetivo de ponderar la adecuación de la temática de cada artículo respecto a los objetivos del presente trabajo. Como resultado inmediato de esta revisión, se eliminaron veinte (20) artículos que claramente no respondían a la temática propuesta. Para el remanente de los trabajos se elaboró una tabla de clasificación de acuerdo con la presencia de los ejes temáticos mencionados en la introducción: *Reinforcement Learning* (RL), *Multi-Agent Systems* (MAS) y *Transfer Learning* (TL). Para los trabajos que presentaban los tres (3) tópicos se asignó una valoración de diez (10) o nueve (9), atribuido discrecionalmente según la profundidad del contenido aportado. En caso de tener sólo dos (2) de los ejes temáticos la valoración oscilaba entre ocho (8) y seis (6), usando el mismo criterio discrecional. En caso de presentar sólo uno de los ejes, el rango de valoración se ubicaba entre cinco (5) y tres (3). Por último y para los trabajos que a pesar de tener referencias o palabras clave relacionadas con los ejes temáticos, no calificaban para este estudio, se definió un rango de dos (2) y (1).

Tabla 7. Tabla de clasificación de los trabajos

CRITERIO	VALORACIÓN
RL+MAS+TL	10-9
RL+MAS o RL+TL	8-6
RL	5-3
Ninguno	2-1

Una vez aplicada la clasificación mencionada, la distribución de los trabajos quedó estructurada de la siguiente forma:

Tabla 8. Clasificación final de los trabajos (cantidad de trabajos por valoración)

EJES TEMÁTICOS	VALORACIÓN									
	10	9	8	7	6	5	4	3	2	1
RL+MAS+TL	9	51								
RL + MAS			20	35	26					
RL + TL			20	8	8					
RL						15	21	13		
MAS						1	3			
TL						1	3	2		
Ninguno									35	52
Totales	9	51	40	43	34	17	27	15	35	52

Como resultado de la clasificación, se tomó el conjunto de los trabajos con valoración diez (10) o nueve (9), resultando en un total objetivo de sesenta (60) trabajos. Al analizar su distribución por año de producción, se observa que el setenta y tres por ciento (73%) de los mismos fueron producidos a partir del año dos mil diecinueve (2019), lo que constata la actualidad del presente estudio. Adicionalmente fue constatado que todos los trabajos seleccionados fueron escritos en inglés, quedando filtrados en otras etapas estudios en portugués y castellano.

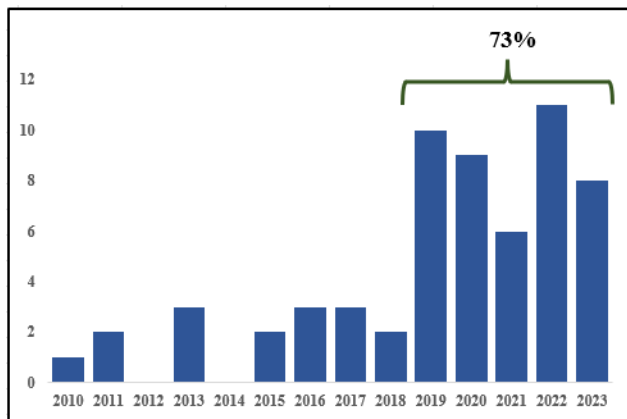


Figura 1. Distribución de los trabajos por año de producción

El proceso completo de selección de los trabajos puede visualizarse en la figura 2.

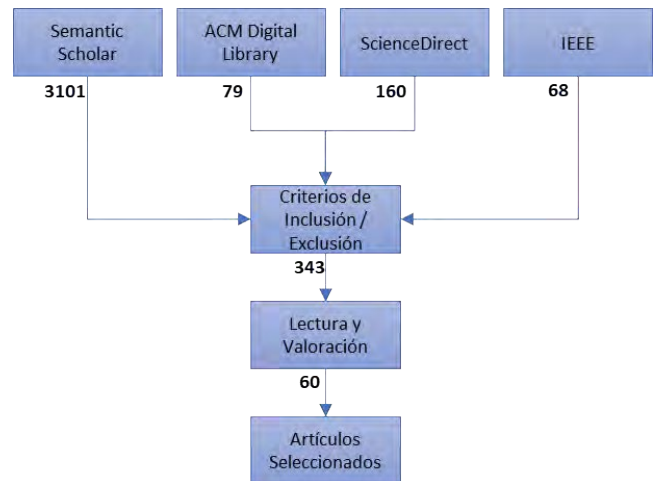


Figura 2. Proceso de selección de trabajos

4. Síntesis de los Datos Extraídos

Mediante la lectura y análisis de los trabajos seleccionados, fue posible preparar la siguiente síntesis de respuestas asociadas a las preguntas de investigación:

P1. ¿Cuáles son los principales métodos de transferencia de conocimiento?

El objetivo de esta pregunta apunta a definir cómo se resolvió el problema de la transferencia de conocimiento en cada trabajo. Durante el proceso de análisis de cada trabajo fue observado que la mayoría de éstos describían métodos específicos para su contexto de aplicación (relacionado con la tercera pregunta del estudio), generando una gran amplitud de resultados. Esta variedad de resultados fue encuadrada estadísticamente según la cantidad de apariciones de los principales métodos, tanto en su forma canónica como en derivadas particulares de cada estudio, que se ve reflejada en la tabla 7.

Tabla 9. Frecuencia de aparición de los métodos para transferencia de conocimiento

MÉTODO	ARTÍCULOS MENCIONADOS
Curriculum learning	12%
Action-value transfer	6%
Teacher-student	6%
Distillation	5%
Federated Learning	4%
Hierarchical Learning	4%
Instance transfer	4%
Model Transfer	4%
Parallel Transfer	4%
Policy Reuse	4%

Otros métodos (30)	47%
--------------------	-----

La tabla 8 lista las apariciones de cada método de transferencia por estudio seleccionado. Si bien la relevancia de cada aparición es distinta (en algunos casos se trata del método aplicado como solución al problema y en otros es solamente parte del análisis), en cuanto a los objetivos de este trabajo resulta relevante cuantificar las menciones independientemente de otras consideraciones.

Tabla 10. Referencias de los métodos de transferencia de conocimiento en la lista de trabajos seleccionados

MÉTODO DE TRANSFERENCIA	ARTÍCULOS
Curriculum learning	[5,9,23,24,39,41,52,55,60]
Action-value transfer	[14,41,42,56,64]
Teacher-student framework	[11,13,26,34,40]
Distillation	[5,9,20,56]
Federated Learning	[30,31,49]
Hierarchical Learning	[7,36,37]
Instance transfer	[14,35,49]
Model Transfer	[35,42,51]
Parallel Transfer	[36,54,56]
Policy Reuse	[13,28,43]
Action Advise	[13,38]
Attention mechanisms	[5,49]
Demonstrations	[9,13]
Feature transfer	[14,49]
Parameter-based transfer	[18,59]
Weight-based Transfer	[30,44]
Adversarial-based domain adaptation	[22]
Adversarial-based Domain (Ad-DISR).	[27]
Behavior transfer	[18]
Bi-directional transfer learning (BDTL)	[19]
Coordinated learning	[21]
Domain adaptation (DA)	[27]
Evolutionary Transfer RL (eTL)	[14]
Experience augmentation	[29]
Function Transfer	[13]
Homomorphism	[17]
Imitation learning	[17]
Learning from Feedback	[13]
Level-adaptive (LA-QTransformer)	[20]
Lifelong federated meta RL (LFMRL)	[49]
Mask-RCNN	[7]
MSDHLA	[6]
Multiagent lateral transfer (MALT)	[39]

Multi-agent transfer soft (MT-SAC).	[48]
Multi-task Learning	[13]
Parameter sharing distributive (CoMP)	[59]
PopulAion-Specific (PASCAL).	[55]
Progressive transfer (PT-DCTL+TS)	[24]
Quaxi	[56]
Transfer via reward shaping	[41]

P2. ¿Cuáles son los factores que impactan en la complejidad del proceso de transferencia del conocimiento?

El objetivo de esta pregunta es identificar los factores de complejidad más comunes dentro del grupo de trabajos seleccionados, definiendo un potencial conjunto de problemas a ser abordados por futuros trabajos. Debido a las variaciones de estilo en la escritura de los artículos y a la falta de consenso respecto a los términos utilizados para definir ciertos factores, fue preciso un análisis semántico de los mismos para generar una lista agrupada de factores, presentada en la tabla 9.

Tabla 11. Resumen de los factores de complejidad encontrados

FACTOR	ARTÍCULOS
Similitud	[7,12,13,14,18,24,27,29,38,39,44,53,63,64]
Transferencia Negativa	[9,13,24,27,44,64]
Definición del objeto de la transferencia	[34,39,51,55]
Simplificación	[18,20,35,60]
Intervención humana	[24,41]
Escalabilidad	[35,36]
Identificación del profesor	[40,51]

El factor más frecuente fue definido en este estudio como “similitud”, y refiere a la premisa de la mayoría de los métodos de transferencia en cuanto a la proximidad estructural de los sujetos y objetos de transferencia. Si bien la mayoría de los estudios exponen experimentos exitosos para demostrar la viabilidad y desempeño de las soluciones, los contextos de demostración parten de la premisa de proximidad entre los agentes y/o las tareas a ser objeto de la transferencia. En el contexto de la transferencia de conocimiento, el término “transferencia negativa” denota eventos que generan un retroceso en el proceso de aprendizaje, generalmente causado por una incorrecta estrategia. En tercer lugar se menciona el aspecto relativo a la propia definición del objeto de aprendizaje. El problema relativo a la simplificación del aprendizaje y su abstracción conceptual, la dependencia de intervención humana, las limitaciones de escala y la

identificación del agente profesor, son los restantes factores definidos.

P3. Desde el punto de vista de la aplicabilidad, ¿en qué contextos se plantea el trabajo, tanto físicos como virtuales?

Los trabajos seleccionados apuntaron diversos enfoques de aplicabilidad de su aporte, los cuales fueron divididos en aplicaciones de alta tecnología (navegación espacial, robótica, autos autónomos, etc.), industrial (redes de comunicaciones, logística, etc.) y académicos/teóricos (sin objeto de aplicación específico). A pesar de la diversidad de objetos listados en los estudios, es destacable que en su amplia mayoría los experimentos sucedieron en simuladores de software, debido a la dificultad de realizar el entrenamiento de los sistemas en un entorno físico, conforme previsto en la introducción.

Tabla 12. Contextos de aplicación de los estudios

CONTEXTO	ARTÍCULOS
Académico / Teórico	[9,10,12,13,14,20,22,25,27,29,30,34,37,38,39,40,41,42,43,44,52,54,56,60,62,64]
Alta Tecnología	[5,6,8,11,15,16,17,21,23,24,28,31,32,36,46,49,50,53,55,57,61]
Mejoras Industriales	[7,18,19,26,33,35,45,47,48,51,58,59,63]

P4. ¿Cuáles son las principales técnicas y herramientas utilizadas?

Si bien existe una gran variedad de algoritmos, técnicas y herramientas en el segmento de *Reinforcement Learning*, a los efectos de representar su impacto en el eje temático definido, fue necesario agrupar los mismos de acuerdo a su origen (cuando son derivados de componentes más genéricos) o a su afinidad (soluciones conceptualmente similares).

Tabla 13. Algoritmos y técnicas mencionadas en los estudios

Algoritmos/Técnicas Mencionados	Referencias
Deep Q-Network (<i>DQN, DRQN, DDQN</i>)	[5,7,12,16,19,28,30,33,35,46,50,56,61,64]
Deep Deterministic Policy Gradients (<i>DDPG, MADDPG, ECR-MADDPG, TS-MADDPG, MATD3, TS-MATD3</i>)	[5,7,11,18,19,25,28,29,35,55,58,59]
Partially observable Markov decision process (<i>POMDP, Dec-POMDP</i>)	[12,20,23,42,43,49,50,55]
Experience replay technique (<i>Hindsight Experience Replay, Prioritized replay buffers</i>)	[7,11,25,29,49]
Monotonic Value Function Factorization (<i>QMLX, NSR-QMLX</i>)	[10,56,59,64]
Proximal Policy Optimization (<i>PPO</i>)	[23,27,31,35]
Trust Region Policy Optimization (<i>TRPO</i>)	[23,30,31,58]
Actor-Critic, Advantage actor-critic (<i>A2C</i>)	[30,31,49]
Centralized training and decentralized execution (<i>CTDE</i>)	[29,49,56]
FCQ-learning, Fully Convolutional Q-Network (<i>FCQN, GCN</i>)	[33,42,61]

CQ-learning	[42,62]
Independent Q-Learning (<i>IQL, IRL</i>)	[10,48]
Long short-term memory (<i>LSTM</i>)	[7,55]
Value decomposition networks (<i>VDN, NSR-VDN</i>)	[56,64]
Bayesian option-based knowledge reuse (<i>Bayes-OKR, Bayes-Pepper, Bayes-ToMoP, BPR+, Deep BPR+</i>)	[9]
Brood War application programming interface (<i>BWAPI</i>)	[25]
Counterfactual Multi-Agent Policy Gradients (<i>COMA</i>)	[10]
Generative Adversarial Network (<i>GAN, CycleGAN, RL-GAN, CURL</i>)	[27]
Decentralized Reinforcement Learning (<i>DEC-RL</i>)	[33]
Distributed Coordinated Learning (<i>DCL</i>)	[21]
DriftER, MDP-CL	[9]
Gaussian mixture model (<i>GMM</i>)	[39]
Genetic encoding	[6]
Graph neural networks (<i>GNN</i>)	[12]
Inductive Graph Reinforcement Learning (<i>IG-RL</i>)	[33]
KT-HA- Q (λ) algorithm	[15]
Maximum margin planning (<i>MMF</i>)	[17]
Networked partially observable stochastic game (<i>ND-POSG</i>)	[59]
NeuroEvolution of augmenting topologies (<i>NEAT</i>)	[6]
Object-Oriented MDPs	[52]
Population Invariant agent with Transformer (<i>PIT</i>)	[20]
Programming apprenticeship learning (<i>LPAL</i>)	[17]
Rainbow	[7]
RL-CD	[9]
Stochastic gradient descent (<i>SGD</i>)	[31]
VAE-Embedding (<i>DARLA, LUSR</i>)	[27]
Value function iteration (<i>VFI</i>)	[6]
Weighted Q-learning	[7]

Conclusiones

El objetivo central del presente trabajo consiste en realizar un mapeo sistemático de los estudios relativos a la transferencia de conocimiento en sistemas multi-agente, entrenados mediante técnicas de aprendizaje por refuerzo. Conforme comentado en la introducción, si bien no se pretendió realizar una continuación del trabajo [4], sirvió éste de base estructural para ordenar la investigación y la adopción de algunas definiciones de nomenclatura provistas. Por otro lado, el presente estudio aporta un enfoque más actualizado, incluyendo un setenta y tres por ciento (73%) de artículos más recientes que [4], además de ampliar la investigación al considerar los factores de complejidad y el contexto de aplicación de cada estudio.

Si bien los hallazgos obtenidos en general responden a las expectativas generadas, y están alineados con trabajos similares anteriores respecto a las soluciones planteadas, se identificaron dos (2) puntos relevantes a ser analizados

más profundamente. El primero de ellos es la falta de adopción de una ontología explícita para los métodos de transferencia de conocimiento en la mayoría de los trabajos abordados, que dificulta el entendimiento de las relaciones entre diferentes algoritmos. Si bien [4] propone una clasificación clara y precisa, la misma no es necesariamente aplicada por otros autores, lo cual se torna en una limitación a la hora de analizar comparativamente distintas propuestas.

El segundo punto a destacar es la prominencia del factor “similitud” en el análisis de los factores que inciden en la complejidad de las soluciones. Aunque este es un factor conocido a priori, la relevancia del mismo se hace evidente por la cantidad de menciones detectada y el impacto que algunos trabajos le asignan, al punto de tener que simplificar fuertemente los ambientes, con el objetivo de llegar a cierta similitud.

A raíz de estas consideraciones se sugiere continuar la investigación a través de dos (2) propuestas investigativas análogas: una revisión profunda y compilación de la ontología existente, que permita clasificar y relacionar los estudios más recientes, y un análisis de las técnicas de simplificación de ambientes, y su impacto en la calidad del aprendizaje.

Referencias

- [1] Uhrmacher, A. M., & Weyns, D. (Eds.). (2009). Multi-Agent systems: Simulation and applications. CRC press.
- [2] SUTTON, R.; BARTO, A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. 344 p.
- [3] Glatt, R., Da Silva, F. L., & Costa, A. H. R. (2016, October). Towards knowledge transfer in deep reinforcement learning. In *2016 5th Brazilian Conference on Intelligent Systems (BRACIS)* (pp. 91-96). IEEE.
- [4] da Silva, F. L., & Costa, A. H. R. (2019). A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*, 64, 645–703.
- [5] Chen, J., Ma, R., & Oyekan, J. (2023). A deep multi-agent reinforcement learning framework for autonomous aerial navigation to grasping points on loads. *Robotics and Autonomous Systems*, 104489.
- [6] Chong, W., Jun, L., Ning, J., Jun, W., & Hao, C. (2011). A distributed cooperative dynamic task planning algorithm for multiple satellites based on multi-agent hybrid learning. *Chinese Journal of Aeronautics*, 24(4), 493–505.
- [7] Rupprecht, T., & Wang, Y. (2022). A survey for deep reinforcement learning in markovian cyber-physical systems: Common problems and solutions. *Neural Networks*, 153, 13–36.
- [8] Tošić, P. T., & Vilalta, R. (2010). A unified framework for reinforcement learning, co-learning and meta-learning how to coordinate in collaborative multi-agent systems. *Procedia Computer Science*, 1(1), 2217–2226.
- [9] Chen, H., Liu, Q., Fu, K., Huang, J., Wang, C., & Gong, J. (2022). Accurate policy detection and efficient knowledge reuse against multi-strategic opponents. *Knowledge-Based Systems*, 242, 108404.
- [10] Hall, G., & Holladay, K. (2020). Adaptive Average Exploration in Multi-Agent Reinforcement Learning. *2020 AIAA/IEEE 39th Digital Avionics Systems Conference (DASC)*, 1–6.
- [11] Ren, H., & Ben-Tzvi, P. (2020). Advising reinforcement learning toward scaling agents in continuous control environments with sparse rewards. *Engineering Applications of Artificial Intelligence*, 90, 103515.
- [12] Chen, L., Chen, Z., Tan, B., Long, S., Gašić, M., & Yu, K. (2019). AgentGraph: Toward universal dialogue management with structured deep reinforcement learning. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 27(9), 1378–1391.
- [13] da Silva, F. L., Warnell, G., Costa, A. H. R., & Stone, P. (2020). Agents teaching agents: a survey on inter-agent transfer learning. *Autonomous Agents and Multi-Agent Systems*, 34, 1–17.
- [14] Hou, Y., Ong, Y.-S., Feng, L., & Zurada, J. M. (2017). An evolutionary transfer reinforcement learning framework for multiagent systems. *IEEE Transactions on Evolutionary Computation*, 21(4), 601–615.
- [15] Ding, D., Ding, Z., Wei, G., & Han, F. (2019). An improved reinforcement learning algorithm based on knowledge transfer and applications in autonomous vehicles. *Neurocomputing*, 361, 243–255.
- [16] Lim, J., Son, H., Lee, D., & Lee, D. (2017). An MARL-based distributed learning scheme for capturing user preferences in a smart environment. *2017 IEEE International Conference on Services Computing (SCC)*, 132–139.
- [17] Boularias, A., & Chaib-Draa, B. (2013). Apprenticeship learning with few examples. *Neurocomputing*, 104, 83–96.
- [18] Wang, T., & Tang, Y. (2022). Automatic voltage control of differential power grids based on transfer learning and deep reinforcement learning. *CSEE Journal of Power and Energy Systems*.
- [19] Sun, G., Wang, X., Jiang, R., Liu, T., & Xu, Y. (2022). Bi-directional Deep Transfer Learning for RIS-enhanced Multi-cell OFDMA Systems. *2022 14th International Conference on Wireless Communications and Signal Processing (WCSP)*, 154–159.
- [20] Zhou, T., Zhang, F., Shao, K., Dai, Z., Li, K., Huang, W., Wang, W., Wang, B., Li, D., Liu, W., & others. (2023). Cooperative Multi-Agent Transfer Learning with Coalition Pattern Decomposition. *IEEE Transactions on Games*.
- [21] Yu, C., Wang, X., & Feng, Z. (2019). Coordinated multiagent reinforcement learning for teams of mobile sensing robots. *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems*, 2297–2299.
- [22] Meng, L., Li, J., & Lu, K. (2022). Cross-domain communications between agents via adversarial-based domain adaptation in reinforcement learning. *ICC 2022-IEEE International Conference on Communications*, 413–418.
- [23] Nguyen, T. T., Hatua, A., & Sung, A. H. (2019). Cumulative training and transfer learning for multi-robots collision-free navigation problems. *2019 IEEE 10th Annual Ubiquitous*

Computing, Electronics & Mobile Communication Conference (UEMCON), 305–311.

- [24] Xia, S.-M., Zhang, L., Bai, W., Zhou, X.-Y., & Pan, Z.-S. (2019). DDoS Traffic Control Using Transfer Learning DQN With Structure Information. *IEEE Access*, 7, 81481–81493.
- [25] Xie, D., & Zhong, X. (2019). Deep Deterministic Policy Gradients with Transfer Learning Framework in StarCraft Micromanagement. *2019 IEEE International Conference on Electro Information Technology (EIT)*, 410–415.
- [26] Yang, H., Xiong, Z., Zhao, J., Niyato, D., Yuen, C., & Deng, R. (2020). Deep reinforcement learning based massive access management for ultra-reliable low-latency communications. *IEEE Transactions on Wireless Communications*, 20(5), 2977–2990.
- [27] Li, D., Meng, L., Li, J., Lu, K., & Yang, Y. (2022). Domain adaptive state representation alignment for reinforcement learning. *Information Sciences*, 609, 1353–1368.
- [28] Spatharis, C., & Blekas, K. (2020). Double deep multiagent reinforcement learning for autonomous driving in traffic maps with road segments and unsignaled intersections. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–6.
- [29] Li, K., Liu, S., & Jia, Q.-S. (2022). Enabling Inter-Agent Transfer for Multi-Agent Learning System by Incorporating Role Reversal. *2022 China Automation Congress (CAC)*, 5667–5672.
- [30] Lim, H.-K., Kim, J.-B., Ullah, I., Heo, J.-S., & Han, Y.-H. (2021). Federated reinforcement learning acceleration method for precise control of multiple devices. *IEEE Access*, 9, 76296–76306.
- [31] Lim, H.-K., Kim, J.-B., Kim, S.-Y., & Han, Y.-H. (2020). Federated reinforcement learning for automatic control in sdn-based iot environments. *2020 International Conference on Information and Communication Technology Convergence (ICTC)*, 1868–1873.
- [32] Prasad, B. K. S., & Ramasangu, H. (2017). Goal-based multi-agent tree formation using reinforcement learning. *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2180–2184.
- [33] Devailly, F.-X., Larocque, D., & Charlin, L. (2021). Ig-rl: Inductive graph reinforcement learning for massive-scale traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), 7496–7507.
- [34] Silva, F. L. da. (2019). Integrating agent advice and previous task solutions in multiagent reinforcement learning. *Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems*, 2447–2448.
- [35] Hu, T., Liao, Q., Liu, Q., & Carle, G. (2023). Inter-Cell Network Slicing with Transfer Learning Empowered Multi-Agent Deep Reinforcement Learning. *IEEE Open Journal of the Communications Society*.
- [36] Kono, H., Murata, Y., Kamimura, A., Tomita, K., & Suzuki, T. (2016). Knowledge Co-creation Framework: Novel Transfer Learning Method in Heterogeneous Multi-agent Systems. *Distributed Autonomous Robotic Systems: The 12th International Symposium*, 389–403.
- [37] Kelly, S., & Heywood, M. I. (2015). Knowledge transfer from keepaway soccer to half-field offense through program symbiosis: Building simple programs for a complex task. *Proceedings of the 2015 Annual Conference on Genetic and Evolutionary Computation*, 1143–1150.
- [38] Markova, V. D., & Shopov, V. K. (2019). Knowledge transfer in reinforcement learning agent. *2019 International Conference on Information Technologies (InfoTech)*, 1–4.
- [39] Shi, H., Li, J., Mao, J., & Hwang, K.-S. (2021). Lateral transfer learning for multiagent reinforcement learning. *IEEE Transactions on Cybernetics*.
- [40] Zhu, C., Cai, Y., Leung, H., & Hu, S. (2020). Learning by reusing previous advice in teacher-student paradigm. *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, 1674–1682.
- [41] Narvekar, S., & Stone, P. (2018). Learning curriculum policies for reinforcement learning. *ArXiv Preprint ArXiv:1812.00285*.
- [42] Hu, Y., Gao, Y., & An, B. (2015). Learning in Multi-agent Systems with Sparse Interactions by Knowledge Transfer and Game Abstraction. *AAMAS*, 753–761.
- [43] Mahmud, M. M. H., & Ramamoorthy, S. (2013). Learning in non-stationary MDPs as transfer learning. *AAMAS*, 1259–1260.
- [44] Song, J., Gao, Y., Wang, H., & An, B. (2016). Measuring the distance between finite Markov decision processes. *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 468–476.
- [45] Huang, B., Chen, X., Sun, Y., & He, W. (2022). Multi-agent cooperative strategy learning method based on transfer Learning. *2022 13th Asian Control Conference (ASCC)*, 1095–1100.
- [46] Kankashvar, M., Bolandi, H., & Mozayani, N. (2023). Multi-agent Q-Learning control of spacecraft formation flying reconfiguration trajectories. *Advances in Space Research*, 71(3), 1627–1643.
- [47] Wu, C., Ma, Z., & Kim, I. (2020). Multi-agent reinforcement learning for traffic signal control: Algorithms and robustness analysis. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1–7.
- [48] Ge, H., Gao, D., Sun, L., Hou, Y., Yu, C., Wang, Y., & Tan, G. (2021). Multi-agent transfer reinforcement learning with multi-view encoder for adaptive traffic signal control. *IEEE Transactions on Intelligent Transportation Systems*, 23(8), 12572–12587.
- [49] Wang, Y., Shang, F., & Lei, J. (2023). Multi-granularity fusion resource allocation algorithm based on dual-attention deep reinforcement learning and lifelong learning architecture in heterogeneous IIoT. *Information Fusion*, 101871.
- [50] Seong, M., Jo, O., & Shin, K. (2023). Multi-UAV trajectory optimizer: A sustainable system for wireless data harvesting with deep reinforcement learning. *Engineering Applications of Artificial Intelligence*, 120, 105891.
- [51] Hu, T., Liao, Q., Liu, Q., & Carle, G. (2022). Network Slicing via Transfer Learning aided Distributed Deep Reinforcement Learning. *GLOBECOM 2022-2022 IEEE Global Communications Conference*, 2909–2914.
- [52] Silva, F. L. da, & Costa, A. H. R. (2018). Object-oriented curriculum generation for reinforcement learning.

Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems, 1026–1034.

- [53] Jiang, Y., Liu, J., & Zheng, H. (2023). Optimal scheduling of distributed hydrogen refueling stations for fuel supply and reserve demand service with evolutionary transfer multi-agent reinforcement learning. *International Journal of Hydrogen Energy*.
- [54] Taylor, A., Dusparic, I., Guériau, M., & Clarke, S. (2019). Parallel transfer learning in multi-agent systems: What, when and how to transfer? *2019 International Joint Conference on Neural Networks (IJCNN)*, 1–8.
- [55] Yan, C., Xiang, X., Wang, C., Li, F., Wang, X., Xu, X., & Shen, L. (2023). PASCAL: PopulAtion-Specific Curriculum-based MADRL for collision-free flocking with large-scale fixed-wing UAV swarms. *Aerospace Science and Technology*, 133, 108091.
- [56] Liang, W., Wang, J., Bao, W., Zhu, X., Wu, G., Zhang, D., & Niu, L. (2022). Qauxi: Cooperative multi-agent reinforcement learning with knowledge transferred from auxiliary task. *Neurocomputing*, 504, 163–173.
- [57] Lee, X. Y., Esfandiari, Y., Tan, K. L., & Sarkar, S. (2021). Query-based targeted action-space adversarial policies on deep reinforcement learning agents. *Proceedings of the ACM/IEEE 12th International Conference on Cyber-Physical Systems*, 87–97.
- [58] Yan, Z., & Wu, C. (2021). Reinforcement learning for mixed autonomy intersections. *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2089–2094.
- [59] Hu, F., Deng, Y., & Aghvami, A. H. (2022). Scalable multi-agent reinforcement learning for dynamic coordinated multipoint clustering. *IEEE Transactions on Communications*, 71(1), 101–114.
- [60] Narvekar, S., Sinapov, J., Leonetti, M., & Stone, P. (2016). Source task creation for curriculum learning. *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, 566–574.
- [61] Schwab, D., Zhu, Y., & Veloso, M. (2020). Tensor action spaces for multi-agent robot transfer learning. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 5380–5386.
- [62] Vrancx, P., de Hauwere, Y.-M., & Nowé, A. (2011). Transfer learning for multi-agent coordination. *International Conference on Agents and Artificial Intelligence*, 2, 263–272.
- [63] Zhao, Q., Jiang, T., Morozs, N., Grace, D., & Clarke, T. (2013). Transfer learning: A paradigm for dynamic spectrum and topology management in flexible architectures. *2013 IEEE 78th Vehicular Technology Conference (VTC Fall)*, 1–5.
- [64] Liu, Y., Hu, Y., Gao, Y., Chen, Y., & Fan, C. (2019). Value Function Transfer for Deep Multi-Agent Reinforcement Learning Based on N-Step Returns. *IJCAI*, 457–463.