

A Systematic Review on Knowledge Transfer in Multi-Agent Systems using Reinforcement Learning

**Marcone de Freitas Marques, Giovani Parente Farias,
Eder Mateus Nunes Gonçalves, Diana Francisca Adamatti**

¹Centro de Ciências Computacionais – Universidade Federal do Rio Grande
Av. Itália, s/n - km 8 - Carreiros – 96.170-000 – Rio Grande – RS – Brazil

{mmsap1998, giovanifarias, dianaada}@gmail.com

edergoncalves@furg.br

Abstract. This paper presents a systematic review on knowledge transfer in multi-agent systems using reinforcement learning. The review followed the PRISMA protocol, analyzing relevant articles from the IEEE Xplore, Scopus, and Web of Science databases. It highlighted research focused on multi-agent reinforcement learning, with works aiming to develop generic algorithms or frameworks for knowledge transfer. Recent approaches enhance agents' efficiency and adaptability, reducing learning time. However, challenges related to communication robustness and knowledge representations' compatibility persist.

Resumo. Este artigo apresenta uma revisão sistemática sobre transferência de conhecimento em sistemas multiagente utilizando aprendizado por reforço. A revisão seguiu o protocolo PRISMA, analisando artigos relevantes nas bases IEEE Xplore, Scopus e Web of Science, revelando pesquisas voltadas para o aprendizado por reforço multiagente, com trabalhos buscando desenvolver algoritmos ou frameworks genéricos para a transferência de conhecimento. Abordagens recentes melhoraram a eficiência e a adaptabilidade dos agentes, reduzindo o tempo de aprendizado. Porém, desafios relacionados à robustez da comunicação e à compatibilidade das representações de conhecimento ainda persistem.

1. Introdução

Sistemas Multiagente (SMA) constituem uma abordagem computacional que envolve múltiplos agentes autônomos capazes de interagir entre si e com o ambiente, com o intuito de alcançar objetivos específicos, sejam eles individuais ou coletivos [Zambonelli et al. 2001, Wooldridge 2009]. Esses sistemas têm se destacado na modelagem de ambientes dinâmicos e distribuídos, sendo aplicados em uma ampla gama de domínios, como robótica colaborativa, gerenciamento de tráfego, sistemas de saúde, redes inteligentes e simulações sociais [Stone and Veloso 2000, Vinyals 2019].

Um dos principais desafios num SMA reside na coordenação de agentes heterogêneos, cujas capacidades cognitivas, níveis de conhecimento e objetivos podem ser distintos. A comunicação eficiente e o compartilhamento de conhecimento tornam-se, assim, fatores críticos para o desempenho global do sistema [Schillo et al. 2001, Panait and Luke 2005]. A capacidade dos agentes de colaborar, adaptar-se e aprender

com outros agentes, em um ambiente não determinístico e possivelmente adversarial, demanda mecanismos robustos de transferência de conhecimento. Agentes podem possuir níveis de conhecimento e habilidades distintos, tornando, desta forma, a colaboração entre os mesmos um fator essencial para a eficiência do sistema. Assim, mecanismos que permitam a transferência de conhecimento entre agentes são fundamentais para um melhor desempenho do SMA.

A transferência de conhecimento num SMA refere-se ao processo pelo qual agentes compartilham experiências, políticas ou modelos de comportamento com o objetivo de acelerar o aprendizado e melhorar a adaptabilidade coletiva [Taylor and Stone 2009, Gao et al. 2021]. Tal processo pode ocorrer em diferentes níveis: (i) no nível individual, em que agentes adquirem conhecimento diretamente de outros por meio de observação ou comunicação; e (ii) no nível sistêmico, onde o conhecimento é compartilhado globalmente para otimizar a performance do sistema como um todo [Kono et al. 2014, Gupta 2017].

Diversas estratégias têm sido propostas para viabilizar essa transferência, incluindo comunicação explícita, aprendizado por imitação, aprendizado baseado em regras, aprendizado por demonstração, e aprendizado por reforço [Lu and Yan 2020, Zhu and Yu 2019]. Cada uma dessas abordagens apresenta vantagens e limitações, dependendo do contexto de aplicação. Por exemplo, enquanto a comunicação explícita oferece agilidade na disseminação de informação, o aprendizado por reforço proporciona maior flexibilidade e adaptabilidade em ambientes dinâmicos.

O aprendizado por reforço tem emergido como uma técnica promissora para a construção de agentes inteligentes em ambientes multiagente, possibilitando o aprendizado baseado na interação com o ambiente e com outros agentes [Busoniu et al. 2008, Long et al. 2024]. No contexto de SMA, o aprendizado por reforço multiagente (*Multi-Agent Reinforcement Learning* – MARL) permite que agentes aprendam políticas cooperativas ou competitivas, dependendo da natureza da tarefa [Zhang et al. 2021]. Além disso, estratégias de transferência de conhecimento aplicadas ao MARL, como transferência de políticas, transferência de recompensas e aprendizado federado, têm se mostrado eficazes na aceleração da convergência do aprendizado e na redução do custo computacional [Bou-Ammar et al. 2015, Zhu 2023].

Apesar dos avanços recentes, persistem desafios relevantes na aplicação de transferência de conhecimento em SMA. Entre eles, destacam-se a compatibilidade entre representações de conhecimento, a robustez das comunicações, a confiabilidade das informações transferidas, e a necessidade de mecanismos que evitem a sobrecarga computacional ou o aprendizado negativo [Silva and Costa 2019]. A garantia de que o conhecimento transferido seja útil e não prejudique a performance do agente destinatário permanece um problema aberto.

Diante da relevância crescente do tema, este artigo apresenta uma revisão sistemática da literatura sobre transferência de conhecimento em SMA com ênfase em abordagens baseadas em aprendizado por reforço. Através de bases de dados acadêmicas consolidadas, são analisadas as contribuições mais recentes, no que se diz respeito ao tema de interesse aqui estabelecido.

2. Metodologia

A revisão sistemática foi conduzida com o objetivo de identificar e analisar estudos sobre transferência de conhecimento em SMA que utilizem aprendizado por reforço. A metodologia seguiu as diretrizes do protocolo PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*) [Moher et al. 2009], apresentada a seguir.

2.1. Questão de Pesquisa

A revisão foi orientada pela seguinte questão de pesquisa: “*Quais os contextos mais atuais onde o aprendizado por reforço vem sendo utilizado na transferência de conhecimento em um contexto multiagente?*”. Para delimitar o escopo da revisão, utilizamos a estratégia PICo (População, Interesse e Contexto):

- **População:** sistemas multiagente.
- **Interesse:** métodos de transferência de conhecimento.
- **Contexto:** aprendizado por reforço.

2.2. Estratégia de Busca

A busca pelos artigos de interesse foi realizada por meio das bases de dados IEEE Xplore, Scopus e Web of Science, com o objetivo de garantir uma ampla cobertura de publicações relevantes. Os termos de busca foram definidos com base na questão de pesquisa e refinados por meio de uma análise iterativa, utilizando tentativa e erro para identificar as melhores combinações de palavras-chave com operadores booleanos. Dessa forma, a chave de pesquisa apresentada a seguir foi aplicada nas bases mencionadas no dia 11 de fevereiro de 2025.

((“*Multi-Agent Reinforcement Learning*”)) AND ((“*Transfer Learning*”))

A busca destas palavras-chave combinadas pelo operador booleano foi realizada no *Abstract* dos artigos.

2.3. Critérios de Inclusão e Exclusão

Para garantir a relevância e adequação dos artigos ao tema proposto, foram definidos critérios objetivos de inclusão e exclusão, aplicados após a execução da busca nas bases selecionadas.

- **Critérios de Inclusão:**
 - Artigos que abordem transferência de conhecimento em sistemas multiagentes.
 - Estudos que utilizem aprendizado por reforço no contexto de interesse.
 - Trabalhos completos publicados.
- **Critérios de Exclusão:**
 - Artigos que não estejam alinhados ao escopo desta pesquisa.
 - Publicações que não estejam redigidas em língua inglesa.
 - Trabalhos publicados antes do ano de 2024.
 - Artigos duplicados entre as bases de dados.
 - Documentos sem acesso ao texto completo.

2.4. Processo de Seleção dos Artigos

Com os critérios previamente definidos, foi realizada a busca sistemática nas bases selecionadas, utilizando a chave de pesquisa construída. A filtragem inicial considerou os critérios de inclusão e exclusão para delimitar o escopo da análise. A Tabela 1 apresenta o número de artigos selecionados em cada base de dados, após a aplicação dos filtros estabelecidos.

Tabela 1. Número de artigos encontrados em diferentes bases de pesquisa.

Base de pesquisa	Número de artigos encontrados
Web of Science	10
Scopus	14
IEEE Xplore	7
Total	31

Em uma etapa subsequente, foi realizada a verificação de duplicidade entre os 31 artigos inicialmente encontrados nas três bases de dados. Identificou-se a existência de 9 duplicatas, que foram removidas, resultando em um total de 22 artigos distintos.

Na sequência, procedeu-se à análise dos títulos dos artigos, sendo identificado um trabalho cujo tema não estava alinhado com os objetivos da pesquisa. Com isso, o número de artigos foi reduzido para 21.

A última etapa do processo de triagem, antes da leitura completa, consistiu na leitura dos resumos dos 21 artigos remanescentes. Com base nos critérios de inclusão e exclusão previamente estabelecidos, foram eliminados 11 artigos nesta fase, totalizando 10 artigos selecionados para leitura e análise aprofundadas.

3. Resultados

A revisão sistemática resultou na seleção de 10 artigos considerados relevantes, os quais abordam a temática de transferência de conhecimento em sistemas multiagentes com o uso de técnicas de aprendizado por reforço. A Tabela 2 apresenta esses trabalhos juntamente com seus principais objetivos.

3.1. Sumarização dos Trabalhos

O estudo de [Mehboob et al. 2024] apresenta uma abordagem que combina transferência de conhecimento e sistemas multiagentes para a parametrização térmica de motores de tração por indução, visando encontrar os melhores valores de temperatura em diferentes componentes do motor. O método baseia-se em um modelo de rede térmica paramétrica. Ao utilizar modelos pré-treinados nessa rede, observa-se uma redução significativa da necessidade de experimentação em laboratório, além de promover uma solução eficaz para lidar com a complexidade dos sistemas dinâmicos. Assim, o trabalho contribui para o desenvolvimento de estratégias de gerenciamento térmico mais adaptativas e eficientes na propulsão ferroviária.

O artigo de [Wang et al. 2024] propõe um framework de tomada de decisão para combate aéreo multiagente, com foco em mecanismos de atenção ao campo de batalha para melhorar a robustez e a transferência em modelos de aprendizado de máquina. A

Tabela 2. Artigos selecionados e seus principais objetivos.

	Trabalho (autor/ano)	Objetivo Principal
1	[Mehboob et al. 2024]	Avaliar o uso de modelos pré-treinados (transferência inicial) para reduzir testes em laboratório e tornar o sistema mais dinâmico.
2	[Wang et al. 2024]	Aplicar aprendizado por reforço profundo em simulações de combate aéreo, comparando agentes que iniciam do zero com aqueles que utilizam transferência de conhecimento.
3	[Chen and Zhang 2024]	Utilizar simetrias euclidianas para identificar similaridades entre tarefas e acelerar a geração de políticas em problemas correlatos.
4	[Kim et al. 2024]	Aprimorar a transferência de conhecimento por meio da randomização de rede em ambientes de aprendizado por reforço multiagente.
5	[Du et al. 2024]	Planejar trajetórias para veículos aéreos não tripulados em missões de resgate, utilizando modelos com transferência de conhecimento na fase inicial.
6	[Siddiqua et al. 2024]	Empregar aprendizado por reforço multiagente co-evolutivo com múltiplas políticas simultâneas, utilizando representações independentes de estados e ações para facilitar a transferência de conhecimento.
7	[Wang and Zhu 2024]	Propor a transferência de políticas online com aprendizado por reforço profundo em ambientes multiagentes, com foco em mecanismos de atenção e abstração temporal.
8	[Namiki et al. 2024]	Implementar uma abordagem que evita a transferência negativa, realizando a transferência apenas de experiências positivas em aprendizado por reforço profundo multiagente.
9	[Liu et al. 2024]	Apresentar um framework com nova arquitetura de rede neural, incorporando aprendizado por transferência com currículo, permitindo a adaptação facilitada a novos contextos.
10	[Zhang et al. 2025]	Desenvolver um algoritmo de aprendizado por reforço multiagente colaborativo para veículos de superfície não tripulados, com transferência de políticas em tempo real e adaptação dinâmica para cercar alvos móveis e evitar obstáculos.

metodologia utiliza aprendizado por reforço profundo, com o treinamento inicial em um ambiente simples (2x2), seguido de um ambiente mais complexo (5x5). Nesses cenários, os agentes reaproveitam habilidades previamente adquiridas. Os resultados indicam um aumento expressivo na taxa de sucesso, na velocidade de convergência e na robustez dos modelos transferidos em comparação a modelos treinados do zero. As conclusões sugerem que essa abordagem pode ser potencialmente validada em ambientes ainda mais complexos.

Em [Chen and Zhang 2024], é investigada a eficiência do aprendizado por reforço multiagente com a exploração de simetrias euclidianas. Para os testes, foram criadas subclasses de jogos de Markov que admitem essas simetrias, e arquiteturas de redes neurais com restrições simétricas foram projetadas como viés indutivo, resultando em desempenho superior em *benchmarks* cooperativos e grande capacidade de generalização. Assim, políticas aprendidas em uma tarefa podem ser eficientemente transferidas para tarefas similares, aumentando a eficiência do aprendizado. As conclusões destacam que a incorporação dessas simetrias não apenas melhora o desempenho em tarefas específicas, mas também oferece uma compreensão mais profunda das estruturas dos problemas de MARL, abrindo novas possibilidades de pesquisa.

O trabalho de [Kim et al. 2024] busca aprimorar a transferência de conhecimento em aprendizado por reforço multiagente por meio da randomização de redes multiagentes. Esse método introduz aleatoriedade nos dados de treinamento ao reiniciar os parâmetros da rede a cada iteração, diversificando o espaço de estados e aumentando a robustez do aprendizado. Os resultados mostram uma melhora de 7,29% na taxa de acerto e maior robustez em relação a métodos tradicionais, evidenciando uma melhor capacidade de generalização. Conclui-se que o MANR (*Multi-Agent Network Randomization*) é uma abordagem promissora para transferência de conhecimento em ambientes complexos, com potencial aplicação em sistemas autônomos e robóticos.

Em [Du et al. 2024], foi desenvolvido um método de planejamento de trajetórias para veículos aéreos não tripulados em cenários de resgate, utilizando aprendizado por reforço profundo aprimorado com redes neurais gráficas. Os resultados indicam que a abordagem acelera o treinamento em até 65% e reduz o número de parâmetros treináveis em 23%. A transferência de conhecimento é realizada ao inicializar o modelo da tarefa-alvo com os parâmetros de uma tarefa-fonte, permitindo que os agentes adaptem suas decisões a novas situações com base em experiências anteriores, otimizando suas operações em ambientes desafiadores.

O estudo de [Siddiqua et al. 2024] desenvolveu um *framework* de aprendizado por reforço multiagente coevolutivo para o treinamento simultâneo de múltiplas políticas na plataforma SMAC (*StarCraft Multi-Agent Challenge*)¹. Os resultados evidenciam melhorias significativas no desempenho, com os agentes aprendendo estratégias eficazes ao compartilhar o conhecimento adquirido em diferentes cenários. A transferência de conhecimento é facilitada por representações independentes de estados e ações, o que permite aos agentes adaptarem suas habilidades e lidarem com a complexidade crescente ao enfrentarem oponentes em constante evolução.

Em [Wang and Zhu 2024], é proposto um algoritmo de transferência de políticas

¹<https://www.techtarget.com/>

online utilizando aprendizado por reforço profundo em ambientes multiagentes. O algoritmo incorpora mecanismos de atenção e abstração temporal como estratégias para aumentar a eficiência do aprendizado. Os resultados indicam uma convergência mais rápida e menor custo computacional em comparação com métodos tradicionais.

O trabalho de [Namiki et al. 2024] apresenta uma abordagem de transferência de conhecimento baseada em aprendizado por reforço profundo multiagente, com foco na prevenção da transferência negativa. A técnica reutiliza parâmetros de múltiplas fontes, reduzindo o tempo de aprendizado e minimizando problemas como neurônios inativos e ações enviesadas. Os resultados demonstram que os agentes se adaptam mais rapidamente a novos ambientes, utilizando conhecimento pré-adquirido para facilitar a exploração e melhorar o desempenho.

Em [Liu et al. 2024], os autores propõem um *framework* de aprendizado por reforço profundo multiagente com transferência de conhecimento, com o objetivo de reduzir o tempo de convergência e a complexidade do treinamento em novos cenários. O *framework* inclui uma nova arquitetura de rede neural e um mecanismo de transferência baseado em currículo. Os experimentos indicam melhorias na eficiência do treinamento e no desempenho final, superando métodos tradicionais e recentes. O estudo conclui que o *framework* facilita a adaptação a novos ambientes sem a necessidade de reestruturações completas da rede, otimizando o tempo e a precisão do treinamento.

Por fim, em [Zhang et al. 2025], foi desenvolvido um algoritmo de aprendizado por reforço multiagente para veículos de superfície não tripulados, com o objetivo de permitir colaboração eficiente na tarefa de cercar um alvo em movimento, evitando obstáculos. Os resultados comprovam a eficácia do método em cenários complexos, com validação experimental em uma piscina interna. A conclusão destaca que a transferência de conhecimento entre os agentes ocorre por meio de um mecanismo de compartilhamento de políticas, permitindo adaptações em tempo real e equilibrando os benefícios individuais e coletivos durante as missões.

3.2. Análise dos Trabalhos Encontrados

As pesquisas identificadas que envolvem transferência de conhecimento em sistemas multiagentes estão majoritariamente centradas na técnica de aprendizado por reforço. Observa-se, a partir desta RSL, que considerou apenas o ano de 2024, um forte predomínio dessa abordagem. Quanto às aplicações, elas são bastante diversas, abrangendo desde a parametrização térmica de motores de tração [Mehboob et al. 2024] até estratégias de tomada de decisão em cenários de combate aéreo [Wang et al. 2024] e de veículos aéreos não tripulados [Du et al. 2024].

Por outro lado, alguns estudos concentraram-se no desenvolvimento de algoritmos e frameworks com maior grau de generalidade, visando modelar aspectos estruturais comuns entre diferentes tarefas. É o caso de [Chen and Zhang 2024], que explora simetrias euclidianas como forma de incorporar conhecimento transferível entre agentes; [Kim et al. 2024], que propõe um método de randomização para diversificar o espaço de estados e aumentar a robustez do aprendizado; e [Wang and Zhu 2024], que utiliza mecanismos de atenção e abstração temporal para extrair e transferir informações relevantes ao contexto da aplicação. Esses trabalhos representam um avanço na área, pois buscam soluções mais generalizáveis, capazes de serem aplicadas a múltiplos domínios sem a

necessidade de reestruturações específicas.

Adicionalmente, destaca-se o trabalho de [Zhang et al. 2025], o único que aborda explicitamente a colaboração entre agentes na construção de uma política conjunta. Tal abordagem vai além das técnicas convencionais de aprendizado por reforço individualizado, ao considerar o compartilhamento de conhecimento como estratégia para melhorar o desempenho coletivo dos agentes em ambientes multiagentes dinâmicos.

4. Conclusão

A revisão sistemática evidenciou que a transferência de conhecimento em sistemas multiagentes, por meio do aprendizado por reforço, constitui uma abordagem promissora para aprimorar a adaptação e o desempenho dos agentes em ambientes dinâmicos. Os estudos analisados demonstram que diversas técnicas contribuem significativamente para a eficiência do aprendizado, reduzindo o tempo de convergência e minimizando a necessidade de treinamento extensivo.

Apesar dos avanços, ainda persistem desafios relevantes, como a robustez na comunicação entre agentes e a compatibilidade entre diferentes representações de conhecimento. Nesse contexto, futuros trabalhos devem se concentrar no desenvolvimento de métodos que reduzam a perda de informação durante o processo de transferência, além de explorar abordagens híbridas que combinem múltiplas técnicas de aprendizado, a fim de potencializar os benefícios dos sistemas multiagentes.

Adicionalmente, embora a transferência de conhecimento por meio do aprendizado por reforço multiagente tenha sido amplamente aplicada em diferentes domínios, outras abordagens também apresentam potencial e merecem maior atenção. Um exemplo é a utilização de estruturas organizacionais entre agentes, conforme proposto em [Farias et al. 2024], que oferece uma perspectiva alternativa e complementar para a transferência de conhecimento. Assim, pesquisas futuras podem expandir significativamente os horizontes desta área, contribuindo para o desenvolvimento de soluções mais robustas, adaptativas e generalizáveis em sistemas multiagentes.

5. Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

Referências

- Bou-Ammar, H., Eaton, E., Luna, J.-M., and Ruvolo, P. (2015). Autonomous cross-domain knowledge transfer in lifelong policy gradient reinforcement learning. In *International Joint Conference on Artificial Intelligence*.
- Busoniu, L., Babuska, R., and De Schutter, B. (2008). A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems, Man, and Cybernetics*, 38(2):156–172.
- Chen, D. and Zhang, Q. (2024). e(3)-equivariant actor-critic methods for cooperative multi-agent reinforcement learning.

- Du, Y., Qi, N., Li, X., Xiao, M., Boulogiorgos, A.-A. A., Tsiftsis, T. A., and Wu, Q. (2024). Distributed multi-uav trajectory planning for downlink transmission: A gnn-enhanced drl approach. *IEEE Wireless Communications Letters*, 13(12):3578–3582.
- Farias, G., Rodrigues, P., Adamatti, D., and Gonçalves, E. (2024). Model for knowledge transfer in agent organizations: a case study on moise+. *The International FLAIRS Conference Proceedings*, 37(1).
- Gao, Z., Xu, K., Ding, B., and Wang, H. (2021). Knowru: Knowledge reuse via knowledge distillation in multi-agent reinforcement learning. *Entropy*, 23(8).
- Gupta, J. K. e. a. (2017). Cooperative multi-agent control using deep reinforcement learning. *International Conference on Autonomous Agents and MultiAgent Systems (AA-MAS)*.
- Kim, D., Jung, H., and Bae, J. (2024). Multi-agent network randomization method for robust knowledge transfer in deep multi-agent reinforcement learning. In *2024 24th International Conference on Control, Automation and Systems (ICCAS)*, pages 325–330.
- Kono, H., Kamimura, A., Tomita, K., Murata, Y., and Suzuki, T. (2014). Transfer learning method using ontology for heterogeneous multi-agent reinforcement learning. *International Journal of Advanced Computer Science and Applications*, 5(10).
- Liu, Y., Wu, X., Bo, Y., Wang, J., and Ma, L. (2024). A transfer learning framework for deep multi-agent reinforcement learning. *IEEE/CAA Journal of Automatica Sinica*, 11(11):2346–2348.
- Long, J., Yu, D., Wen, G., Li, L., Wang, Z., and Chen, C. L. P. (2024). Game-based backstepping design for strict-feedback nonlinear multi-agent systems based on reinforcement learning. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):817–830.
- Lu, Y. and Yan, K. (2020). Algorithms in multi-agent systems: A holistic perspective from reinforcement learning and game theory. *ArXiv*, abs/2001.06487.
- Mehboob, F., Fattouh, A., and Sahoo, S. (2024). Synergizing transfer learning and multi-agent systems for thermal parametrization in induction traction motors. *Applied Sciences*, 14(11).
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group*, t. (2009). Preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Annals of internal medicine*, 151(4):264–269.
- Namiki, M., Okawa, Y., Morita, N., Kakuta, J., and Ogawa, M. (2024). Transfer learning with less negative transfer for multi-agent reinforcement learning: Application and evaluation in base station control. In *2024 IEEE 35th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–7.
- Panait, L. and Luke, S. (2005). Cooperative multi-agent learning: The state of the art. *Autonomous agents and multi-agent systems*, 11(3):387–434.
- Schillo, M., Bürckert, H.-J., Fischer, K., and Klusch, M. (2001). Towards a definition of robustness for market-style open multi-agent systems. In *Proceedings of the Fifth*

- International Conference on Autonomous Agents*, AGENTS '01, page 75–76, New York, NY, USA. Association for Computing Machinery.
- Siddiqua, A., Liu, S., Siddika Nipu, A., Harris, A., and Liu, Y. (2024). Co-evolving multi-agent transfer reinforcement learning via scenario independent representation. *IEEE Access*, 12:99439–99451.
- Silva, F. and Costa, A. (2019). A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*, 64.
- Stone, P. and Veloso, M. (2000). Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots*, 8(3):345–383.
- Taylor, M. E. and Stone, P. (2009). Transfer learning for reinforcement learning domains: A survey. *Journal of Machine Learning Research*, 10:1633–1685.
- Vinyals, O. e. a. (2019). Grandmaster level in starcraft ii using multi-agent reinforcement learning. *Nature*, 575:350–354.
- Wang, C. and Zhu, X. (2024). Ata-maopt: Multi-agent online policy transfer using attention mechanism with time abstraction. *IEEE Access*, 12:158282–158291.
- Wang, Y., Ma, Y., Wang, J., Yuan, H., Wang, M., and Ma, H. (EasyChair, 2024). Multi-agent air combat decision-making based on battlefield attention information. Easy-Chair Preprint 15393.
- Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. Wiley Publishing, 2nd edition.
- Zambonelli, F., Jennings, N. R., and Wooldridge, M. (2001). Organisational rules as an abstraction for the analysis and design of multi-agent systems. *International Journal of Software Engineering and Knowledge Engineering*, 11(03):303–328.
- Zhang, C., Zeng, R., Lin, B., Zhang, Y., Xie, W., and Zhang, W. (2025). Multi-usv cooperative target encirclement through learning-based distributed transferable policy and experimental validation. *Ocean Engineering*, 318:120124.
- Zhang, K., Yang, Z., and Başar, T. (2021). Multi-agent reinforcement learning: A selective overview of theories and algorithms. *Handbook of Reinforcement Learning and Control*.
- Zhu, H. e. a. (2023). A survey of knowledge transfer for deep reinforcement learning in multi-agent systems. *Neurocomputing*, 522:1–21.
- Zhu, T. and Yu, P. S. (2019). Applying differential privacy mechanism in artificial intelligence. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, pages 1601–1609.