SCIENCE CHINA
Information Sciences

• LETTER •

July 2026, Vol. 69, Iss. 7, 179101:1–179101:2
https://doi.org/10.1007/s11432-025-4724-0

LaTac: latent tactics for robust multi-agent coordination under intermittent communication

Enguang YAO[1], Qidong LIU[1,2,3*], Jiajia HOU[1], Mingfei SUN[4], Zongzhang ZHANG[5] & Mingliang XU[1,2,3*]

[1]*School of Computer and Artificial Intelligence, Zhengzhou University, Zhengzhou 450001, China*
[2]*National Supercomputing Center in Zhengzhou, Zhengzhou 450001, China*
[3]*Engineering Research Center of Intelligent Swarm Systems, Ministry of Education, Zhengzhou 450001, China*
[4]*Department of Computer Science, The University of Manchester, Manchester M139PL, UK*
[5]*National Key Laboratory for Novel Software Technology, Nanjing University, Nanjing 210023, China*

Received 9 July 2025/Revised 23 October 2025/Accepted 16 December 2025/Published online 8 June 2026

Citation Yao E G, Liu Q D, Hou J J, et al. LaTac: latent tactics for robust multi-agent coordination under intermittent communication. Sci China Inf Sci, 2026, 69(7): 179101, https://doi.org/10.1007/s11432-025-4724-0

Multi-agent reinforcement learning (MARL) has emerged as a powerful framework for addressing complex challenges across diverse domains [1], including robotics, game AI, and social networks. In most real-world scenarios, agents face partial observability and lack a comprehensive overview to access global information. Consequently, communication becomes essential for agents to share knowledge, sparking increased interest among researchers in MARL enhanced with communication mechanisms (Comm-MARL) [2].

Previous studies have utilized advanced techniques, such as graph neural networks and attention mechanisms, to devise sophisticated communication protocols that process information from different teammates at each timestep. Although these methods achieve strong performance under the assumption of stable and reliable communication, such an assumption rarely holds in realistic multi-agent environments, where teammate information becomes intermittently available, leading to unpredictable gaps in coordination. This presents a systemic challenge analogous to the fog of war in military decision-making, where uncertainty and degraded information flow disrupt coordinated actions. We term this phenomenon intermittent communication, emphasizing not merely constrained capacity, but the temporal unpredictability and unreliability in the availability of teammate information. This focus departs from prior assumptions of stable or selectively scheduled communication by introducing temporally unpredictable disruptions that break the continuity of teammate information flow.

Inspired by competitive games such as football, where coaches tailor tactics (i.e., assign specific roles to each player) based on players' strengths and opponents' weaknesses, we propose an implicit coordinated decision-making framework based on "latent tactics (LaTac)". Here, latent tactics refer to the intricate and diverse local coordination patterns employed among agent roles in multi-agent coordination tasks. However, in multi-agent systems with intermittent communication, agents often operate with fragmented information, severely restricting their consistent coordination and disrupting the formation of tactical behaviors. In contrast, stable communication enables the emergence of richer latent tactics, as agents can rely on complete contextual information to coordinate effectively. To exploit this asymmetry, and inspired by teacher-student coordination paradigms in large-small language model coordination, we propose a two-stage training framework grounded in knowledge distillation. A teacher model trained under stable communication conditions acquires high-quality latent tactics, which are subsequently distilled into a counterpart operating under intermittent communication. By aligning intermediate representations via structured distillation, LaTac facilitates the transfer of robust coordination behaviors, enabling consistent tactical behavior despite severe communication disruptions.

Method. As illustrated in Figure 1(a), the proposed framework LaTac consists of two core components. (1) Coordinative decision-making model and latent tactics training in stable communication environment. In this phase, we first train a coordinative framework based on latent tactics under ideal communication conditions with full inter-agent information availability, aiming to learn reliable role assignment and joint policy behaviors. Meanwhile, a role feature set $\mathbf{Z}$ is simultaneously constructed and stored during training, serving as a semantic prior for role inference in subsequent deployment. (2) Knowledge distillation for intermittent communication environment. Using knowledge distillation techniques, the tactical decision-making strategies learned under ideal communication conditions are effectively transferred to intermittent communication scenarios, ensuring that the model maintains robust coordinative performance even under communication constraints. During the inference phase, the system employs the optimized robust model for intermittent communication to make tactical decisions and retrieves pre-stored role representations from the feature set $\mathbf{Z}$ to generate reliable decisions. Please refer to Appendix A for more detailed methods.

Experiment. We employ the StarCraft II micromanagement (SMAC) benchmark [3] and multi-agent particle environment (MPE) [4] as our testbeds, and evaluate the performance of LaTac against several state-of-the-art baselines on them. To simulate in-

* Corresponding author (email: ieqdliu@zzu.edu.cn, iexumingliang@zzu.edu.cn)

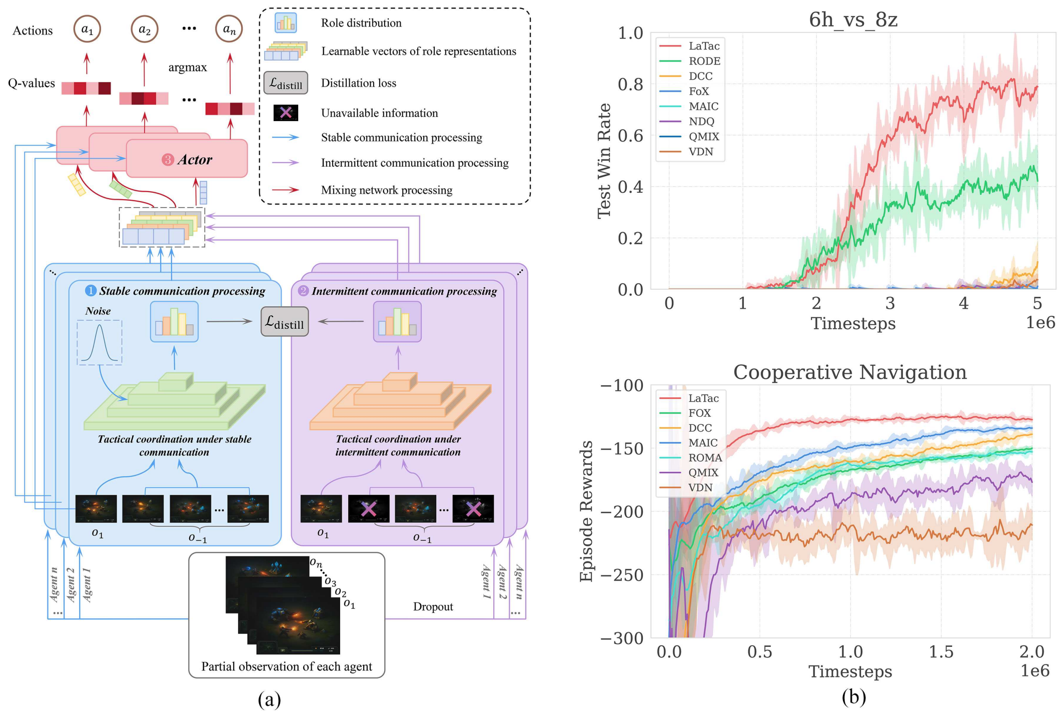

Figure 1 (Color online) (a) Overview of the proposed LaTac framework across communication conditions. The framework adopts a two-phase training paradigm grounded in latent tactics modeling. In the first phase, agents are trained under stable communication to acquire robust coordination behaviors through decentralized information aggregation and tactic-driven decision making (i.e., combining Modules ① and ③). In the second phase, the same modular architecture is trained under intermittent communication, where internal representations—such as aggregated messages and inferred role semantics—are distilled from the stable communication phase (i.e., Module ② learns from the distilled knowledge of Module ①). During execution, Modules ② and ③ are jointly deployed to maintain tactical coherence and robust coordination under intermittent communication. (b) Performance comparison between LaTac and baselines on the SMAC and MPE benchmarks.

termittent communication, we set the information dropout probability to 40%, causing 40% of communications to be randomly dropped.

As shown in Figure 1(b), LaTac consistently outperforms representative baselines across both the SMAC and MPE benchmarks. These results collectively suggest that LaTac's tactic-level abstraction effectively captures latent cooperative structures, enabling resilient multi-agent coordination despite intermittent information flow. Additional experimental results, including ablation and sensitivity analyses, are provided in Appendix B for completeness.

Conclusion. In this study, we analyze the challenges posed by intermittent communication to existing models, and propose a novel multi-agent reinforcement learning framework based on latent tactics, termed LaTac, which effectively mitigates the issue of incomplete information and enhances decision-making accuracy in intermittent communication scenarios. However, LaTac still has room for improvement, such as the difficulty in transferring role embeddings (learned representations of agent roles) trained in one scenario to another. Given the versatility of in large language models (LLMs) in artificial intelligence, which is increasingly being leveraged in reinforcement learning for tasks such as designing reward functions [5], fine-tuning strategies, and improving low-level control, we plan to explore the use of LLMs in the future to generate new roles based on scenarios, rather than being limited to pre-trained roles, enabling reinforcement learning to better generalize to more open and complex environments.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. 62276238, U24A20326, 62325602, 62036010), Foundation of Henan Educational Committee (Grant No. 25HASTIT034), and Natural Science Foundation of Henan (Grant No. 232300421095).

Supporting information Appendixes A–C. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Liu Q, Lian J, Liu C, et al. Enhancing generalization in large-scale HCVRP: a rank-augmented neural solver. In: Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2025. 1845–1856

2 Ding S, Du W, Ding L, et al. Robust multi-agent communication with graph information bottleneck optimization. IEEE Trans Pattern Anal Mach Intell, 2023, 46: 3096–3107

3 Samvelyan M, Rashid T, de Witt C S, et al. The starcraft multi-agent challenge. In: Proceedings of International Conference on Autonomous Agents and MultiAgent Systems, 2019. 2186–2188

4 Mordatch I, Abbeel P. Emergence of grounded compositional language in multi-agent populations. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2018. 1495–1502

5 Qu Y, Jiang Y H, Wang B Y, et al. Latent reward: LLM-empowered credit assignment in episodic reinforcement learning. In: Proceedings of the AAAI Conference on Artificial Intelligence, 2025. 20095–20103