LLMs for Next-Generation Networks: An AI-Native Communication Protocol Perspective

Renxuan Tan

Networked Intelligence for Comprehensive Efficiency (NICE) Lab
College of Information Science and Electronic Engineering
Zhejiang University
http://nice.rongpeng.info/

Aug. 18, 2025

Content

Motivation: Rethinking Protocol Design

Fundamentals of Communication Protocol

- **Definition-** A ruleset governing inter-node message exchange with predetermined formats and temporal sequences.
- **Key Roles-** Ensuring reliability & interoperability, governing timing and sequencing, impacting network key performance indicators (KPIs).
- **Characteristics-** Fixed and standardized. Traditional protocol design workflow is a process that relies on human experts and is both time consuming and exhibits slow iteration cycles.

		Channel Access Schemes					
		TDMA	***frequency division multiple access (FDMA)***			***CDMA***	***Spatial***
			Time-invariant	***Time-variant***			
				Non-OFDMA	***OFDMA***		
MAC protocols	***Contention based***	Aloha CSMA Cognitive Radio	*N.A.*	Cognitive Radio			
	Coordinated	GSM Token Ring Bluetooth	FM/AM radio DVB-T	POTS	LTE 5G New Radio (5GNR)	3G GPS	Satcoms MIMO

Figure: Example of MAC protocols.

Protocol Bottlenecks in Modern Wireless Networks

- General-purpose solutions ⟶ Tailored to specific networks[1]
- Ever-increasing network heterogeneity and scalability[2]
- Rigidity of the structure[3]
- High cost of standardization[4]

Traditional, human-centered protocol design approaches are reaching the upper limits of their capabilities in the face of increasing network complexity, dynamics, and scale.

---

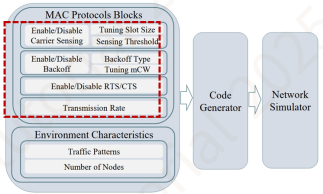
[1] M. P. Mota et al., "The emergence of wireless MAC protocols with multi-agent reinforcement learning," in *IEEE GLOBECOM*, Madrid, Spain, 2021.

[2] S. Sarbu et al., "On scaling latency-aware MAC communication protocols with a hierarchical network topology," in *ICC 2024*, 2024, pp. 2555–2560. DOI: 10.1109/ICC51166.2024.10623079 Accessed: Nov. 8, 2024.

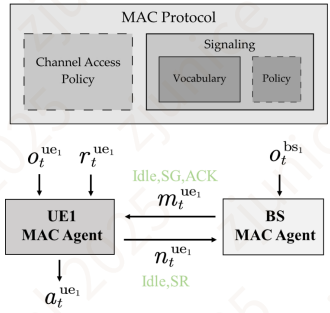
[3] J. Wang et al., *Next-generation Wi-Fi networks with generative AI: Design and insights*, 2024. arXiv: 2408.04835 [cs.NI]. [Online]. Available: https://arxiv.org/abs/2408.04835

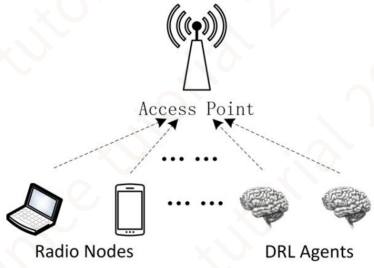
[4] A. Valcarce and J. Hoydis, "Toward joint learning of optimal MAC signaling and wireless channel access," *IEEE Trans. Cognit. Commun. Networking*, vol. 7, no. 4, pp. 1233–1243, 2021 .

DRL4Protocol Applications[5][6][7]



	Action parameter	Values
a_1	Carrier sensing	On or Off
a_2	Slottime	$\{5, 9, 20\}$
a_3	Backoff type	Off, EDID, BEB, Constant
a_4	Minimum CW	$\{15, 31, 63\}$
a_5	RTS/CTS	On or Off
a_7	Datarate	$\{6.5, 13, 19.5, 26, 39, 52, 58.5, 65, 78\} Mbps$





Parameters Configuration **Signaling Learning** **Heterogeneous Multiple Access**

[5] N. Keshtiarast et al., "Wireless MAC protocol synthesis and optimization with multi-agent distributed reinforcement learning," *IEEE Networking Lett.*, vol. 6, no. 4, pp. 242–246, 2024 .

[6] M. P. Mota et al., "Scalable joint learning of wireless multiple-access policies and their signaling," in *IEEE VTC*, Helsinki, Finland, 2022.

[7] Y. Yu et al., "Deep-reinforcement learning multiple access for heterogeneous wireless networks," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7. DOI: 10.1109/ICC.2018.8422168

DRL4Protocol Challenges

- **Generalization**
  - Over-parameterized, fail to generalize outside of their training distributions.
  - Unable to adapt quickly and requires re-training in dynamic environments.
- **Complexity**
  - Selecting the appropriate neural network (NN) architecture and setting hyperparameters are crucial for achieving desired performance levels, requiring domain expertise.
  - RL models are data hungry.
- **Interpretability**
  - NN is also a black-box function, making the protocol operations uninterpretable.
  - Lack of flexibility to control/change rules.
- Efficiency/Scalability/Feasibility

Large Language Model is a promising solution !

Toward Agent and Agentic AI[8]

Stage 1 Generative AI

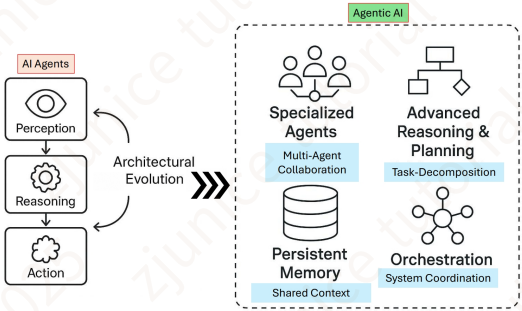
- e.g., LLMs for writing, translating, and chatting
- Prompt ⟼ LLM ⟼ Output
- Direct response

Stage 2 AI Agent

- Autonomous software programs that perform specific tasks
- Prompt ⟼ Tool Call ⟼ LLM ⟼ Output
- Perceive-Reason-Act-Observe Loop

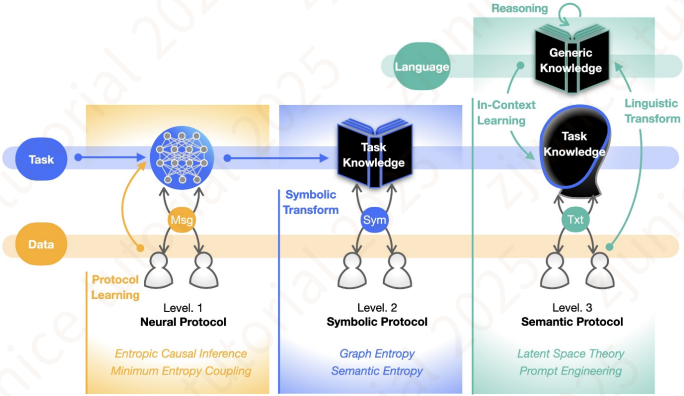
Stage 3 Agentic AI

- Systems of multiple AI agents collaborating to achieve complex goals.
- Goal ⟼ Agent Orchestration ⟼ Output
- Agentic Workflow (Autonomy/Goal-Directed/Dynamic Adaptation/Interactivity)



[8] R. Sapkota et al., *Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges*, 2025. arXiv: 2505.10468 [cs.AI]. [Online]. Available: https://arxiv.org/abs/2505.10468

Three-layer communication protocol evolution[9]



- Roadmap

L1 **Task-oriented** neural protocols

L2 **NN-oriented** symbolic protocols

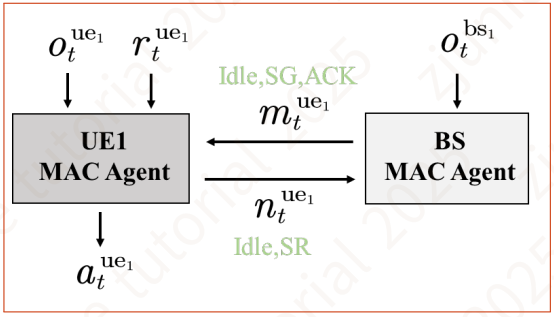
L3 **Language-oriented** semantic protocols

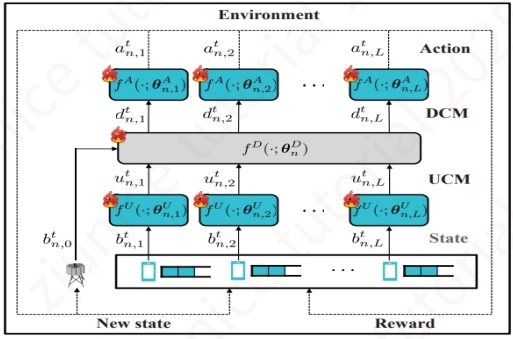
Figure: Three-level categorization of data-driven MAC protocols.

[9] J. Park et al., *Towards semantic communication protocols for 6G: From protocol learning to language-oriented approaches*, 2023. arXiv: 2310.09506 [cs.IT]. [Online]. Available: https://arxiv.org/abs/2310.09506

LLM empowered communication protocols

Still in RL Paradigm – LLM4MAC





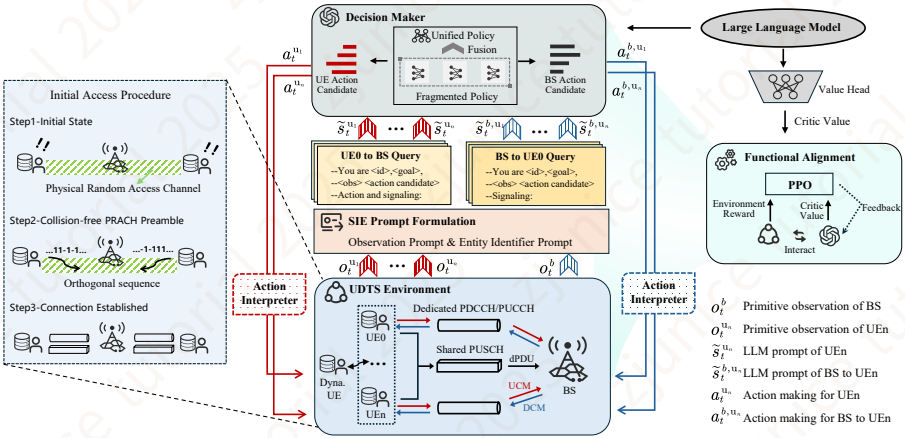
- o_t^{u}: Observation received by the u^{th} UE at time step t. → 缓存的SDU数量。整数，取值：0,...,B
- o_t^{b}: Observation received by the BS at time step t. → 数据信道的状况。整数，取值：0 (信道空)，1...L (仅有UE-i发送)，L+1 (多个UE发送)
- n_t^{u}: The UCM sent from the u^{th} UE at time step t.
- m_t^{u}: The DCM sent to the u^{th} UE at time step t. → bit序列，8bit
- a_t^{u}: Environment action of the u^{th} UE at time step t. → UE动作。整数，取值：0 (什么都不做)，1 (传输最老的SDU)，2 (删除最老的SDU)
- x_t^{u}: Agent internal state of the u^{th} UE at time step t.
- x_t^{b}: Agent internal state of the BS at time step t. → 内部状态，包含最近k个观测、动作、消息 $x_t^{\mathrm{u}} = (o_t^{\mathrm{u}}, \ldots, o_{t-k}^{\mathrm{u}}, a_t^{\mathrm{u}}, \ldots, a_{t-k}^{\mathrm{u}}, n_t^{\mathrm{u}}, \ldots, n_{t-k}^{\mathrm{u}}, m_t^{\mathrm{u}}, \ldots, m_{t-k}^{\mathrm{u}})$

$$x_t^{\mathrm{b}} = (o_t^{\mathrm{b}}, \ldots, o_{t-k}^{\mathrm{b}}, \mathbf{n}_t, \ldots, \mathbf{n}_{t-k}, \mathbf{m}_t, \ldots, \mathbf{m}_{t-k})$$

$$r_t = \begin{cases} +\rho, & \text{if a new SDU was received by the BS} \\ -\rho, & \text{if an UE deleted a SDU that has not been received by the BS} \\ -1, & \text{else,} \end{cases}$$

→ 每一步的reward

Still in RL Paradigm – LLM4MAC[10]



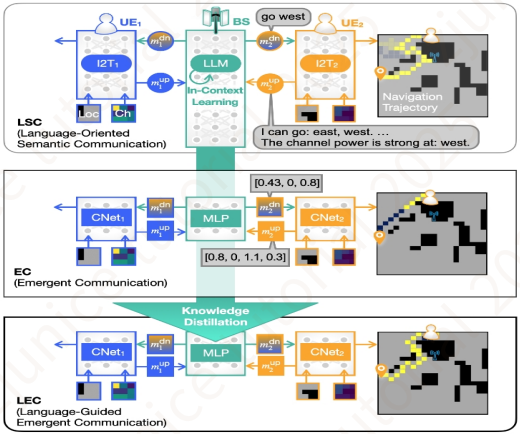
$$\mathbb{P}_{LLM}(a_i \mid p) = \prod_{j=0}^{|a_i|} \mathbb{P}_{LLM}(w_j \mid p, w_{<j}), \mathbb{P}(a_i \mid p) = \frac{e^{\mathbb{LP}_{LLM}(a_i|p)}}{\sum_{a_j \in \mathcal{A}} e^{\mathbb{P}_{LLM}(a_j|p)}}$$

[10] R. Tan et al., *Llm4mac: An llm-driven reinforcement learning framework for mac protocol emergence*, 2025. arXiv: 2503.08123 [cs.NI]. [Online]. Available: https://arxiv.org/abs/2503.08123

Shortcomings of "LLM under the RL paradigm"

- Low data efficiency
  - "Cold Start" problem necessitates extensive, often inefficient exploration for data collection during initial training stages.
  - It demands extensive online environment interactions and meticulous reward function design.
- Low training instability
  - At the cost of thousands of training steps due to the large LLM model size.
  - The training process exhibits sensitivity to hyperparameters, often leading to non-convergence and instability.
- **Fails to capture the essence of LLMs**
  - in-context understanding/reflective reasoning
  - zero-shot operation
  - scaling law

In-Context Learning Paradigm – LLM-enhanced Communication Emergence[11]



- **Multi-Agent Remote Navigation:**
  - Minimizing the steps required to pre-determined destination.
  - Avoiding the weak channels locations.
- **Emergent Communication:**
  - BS: $\boldsymbol{m}^{\mathrm{dn}} = f(m_1^{\mathrm{up}}, \cdots, m_j^{\mathrm{up}}; \theta^{\mathrm{BS}})$
  - UE: $(a_j, m_j^{\mathrm{up}}) = f(s_j, \tilde{m}_j^{\mathrm{dn}}; \theta^{\mathrm{UE}})$
- **Language-Oriented Semantic Communication:**
  - $a_j = f(m_{j,t}^{\mathrm{up}}, \tilde{\boldsymbol{m}}_{j,t-1}^{\mathrm{up}}, \tilde{\boldsymbol{m}}_{j,t-1}^{\mathrm{dn}}, c_K, x'; \theta^{\mathrm{LLM}})$
  - meta instruction x', K-pair examples c_K
- Inject KLD loss between EC and LSC policies

$$\mathcal{L}_{\mathrm{KLD}} := D_{\mathrm{KLD}}\left(p\left(a_j^t \mid \theta\right) \| p\left(a_j^t \mid H^\star\right)\right)$$

$$\theta_{\mathrm{LEC}}^\star = \arg\min_\theta \left\{ \mathbb{E}\left[\sum_{j,t} \left[\left\{ \tilde{r}_j^t + \gamma \max_{a'^t_j}{}^j Q'\left(s_j'^t, a_j'^t \mid \theta\right) - Q\left(s_j^t, a_j^t \mid \theta\right) \right\}^2 + \lambda \mathcal{L}_{\mathrm{KLD}} \right] \right] \right\}$$

[11] Y. Kim et al., *Knowledge distillation from language-oriented to emergent communication for multi-agent remote control*, 2024. arXiv: 2401.12624 [cs.AI]. [Online]. Available: https://arxiv.org/abs/2401.12624

UE 1 : "I can go to : south, east, southeast. From the location (x=0, y=0) aiming for (x=9, y=9), the direction that goes closest to the goal is southeast. Other directions like northeast, east, southwest and south will also get me closer. The channel power is strong at south, east, southeast."
UE 2 : "I can go to : west, southwest, south, east, southeast. From the location (x=0, y=9) aiming for (x=10, y=0), the direction that goes closest to the goal is southwest. Other directions like northwest, west, south and southeast will also get me closer. The channel power is strong at west, south, east, southeast. The channel power is weak at southwest, so the communication service is unstable."
Question : Which direction should UE 1 and UE 2 go?

(a) UL messages $\tilde{m}^{\mathbf{up},t}$ at SNR = 20dB.

UE 1 : " I can go to : south, east, southeast. From the location (x=0, y=0) aiming for (x=9, y=9), the direction that goes closest to the goal is southeast. Other directions like northeast, east, southwest and south will also get me closer. The channel power is strong at south, east, southeast. "
UE 2 : " I can go to : west, southwest, south, east, southeast. From the location (x=0, y=9) aiming for (x=10, y=0), the direction that goes closest to the goal is southwest. Other directions like northwest, west, south, and southeast will also get me closer. The channel power is strong at west, south, east, southeast. The channel power is weak at southwest, so the communication service is unstable. "

(b) UL messages $\tilde{m}^{\mathbf{up},t}$ at SNR = 15dB.

Explanation : UE 1 should choose "southeast" as it brings them closer to the goal and has strong channel power. UE 2 should choose "west" as it brings them closer to the goal, avoiding "southwest" due to weak channel power.
UE 1 : southeast
UE 2 : west

(c) DL message $m^{\mathbf{dn},t}$ at SNR = 20dB, $K = 2$.

Explanation : UE 1 should choose southeast as it brings them closer to the goal and has strong channel power, while avoiding weak channels. UE 2 should choose southwest as it brings them closer to the goal, although southeast is a closer direction. UE 2 should avoid weak channels as much as possible.
UE 1 : southeast
UE 2 : southwest

(d) DL message $m^{\mathbf{dn},t}$ at SNR = 15dB, $K = 2$.

Explanation : What is the goal of the task? What are the constraints of the task? What are the wireless channel strengths?
UE 1 : none
UE 2 : none

(e) DL message $m^{\mathbf{dn},t}$ at SNR = 15dB, $K = 0$.

LLM as Optimizers[12]

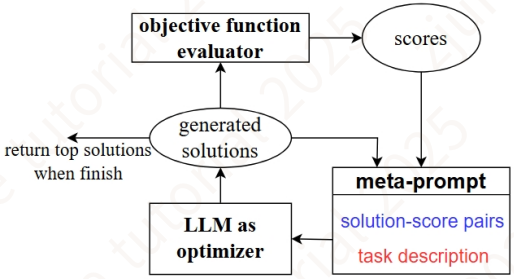


Figure: Optimize by PROmpt (OPRO) framework

The framework is inherently suited to *prompt optimization*, since that problem is itself a well-defined text-optimization task.

Now you will help me minimize a function with two input variables w, b. I have some (w, b) pairs and the function values at those points. The pairs are arranged in descending order based on their function values, where lower values are better.

input:
w=18, b=15
value:
10386334

input:
w=17, b=18
value:
9204724

Give me a new (w, b) pair that is different from all pairs above, and has a function value lower than any of the above. Do not write code. The output must end with a pair [w, b], where w and b are numerical values.

You are given a list of points with coordinates below: (0): (-4, 5), (1): (17, 76), (2): (-9, 0), (3): (-31, -86), (4): (53, -35), (5): (26, 91), (6): (65, -33), (7): (26, 86), (8): (-13, -70), (9): (13, 79), (10): (-73, -86), (11): (-45, 93), (12): (74, 24), (13): (67, -42), (14): (87, 51), (15): (83, 94), (16): (-7, 52), (17): (-89, 47), (18): (0, -38), (19): (61, 58).
Below are some previous traces and their lengths. The traces are arranged in descending order based on their lengths, where lower values are better.

<trace> 0,13,3,16,19,2,17,5,4,7,18,8,1,9,6,14,11,15,10,12 </trace>
length:
2254

<trace> 0,18,4,11,9,7,14,17,12,15,10,5,19,3,13,16,1,6,8,2 </trace>
length:
2017

<trace> 0,11,4,13,6,10,8,17,12,15,3,5,19,2,1,18,14,7,16,9 </trace>
length:
1953

<trace> 0,10,4,18,6,8,7,16,14,11,2,15,9,1,5,19,13,12,17,3 </trace>
length:
1840

Give me a new trace that is different from all traces above, and has a length lower than any of the above. The trace should traverse all points exactly once. The trace should start with <trace> and end with </trace>

LLM as Optimizers 2

Your task is to generate the instruction <INS>. Below are some previous instructions with their scores. The score ranges from 0 to 100.

text:
Let's figure it out!
score:
61

text:
Let's solve the problem.
score:
63

(. . . more instructions and scores . . .)

Below are some problems.

Problem:
Q: Alannah, Beatrix, and Queen are preparing for the new school year and have been given books by their parents. Alannah has 20 more books than Beatrix. Queen has 1/5 times more books than Alannah. If Beatrix has 30 books, how many books do the three have together?
A: <INS>

Ground truth answer:
140

(. . . more exemplars . . .)

Generate an instruction that is different from all the instructions <INS> above, and has a higher score than all the instructions <INS> above. The instruction should begin with <INS> and end with </INS>. The instruction should be concise, effective, and generally applicable to all problems above.

Table 6: Transferability across datasets: accuracies of top instructions found for GSM8K on MultiArith and AQuA.

Scorer	Source	Instruction position	Instruction	Accuracy	
				MultiArith	AQuA
Baselines					
PaLM 2-L	(Kojima et al., 2022)	A_begin	Let's think step by step.	85.7	44.9
PaLM 2-L	(Zhou et al., 2022b)	A_begin	Let's work this out in a step by step way to be sure we have the right answer.	72.8	48.4
PaLM 2-L		A_begin	Let's solve the problem.	87.5	44.1
PaLM 2-L		A_begin	(empty string)	69.3	37.8
text-bison	(Kojima et al., 2022)	Q_begin	Let's think step by step.	92.5	31.9
text-bison	(Zhou et al., 2022b)	Q_begin	Let's work this out in a step by step way to be sure we have the right answer.	93.7	32.3
text-bison		Q_begin	Let's solve the problem.	85.5	29.9
text-bison		Q_begin	(empty string)	82.2	33.5
Ours					
PaLM 2-L	PaLM 2-L-IT on GSM8K	A_begin	Take a deep breath and work on this problem step-by-step.	**95.3**	**54.3**
text-bison	PaLM 2-L-IT on GSM8K	Q_begin	Let's work together to solve math word problems! First, we will read and discuss the problem together to make sure we understand it. Then, we will work together to find the solution. I will give you hints and help you work through the problem if you get stuck.	**96.8**	**37.8**

Figure: Simulation Results

Prompt Engineering Paradigm – Semantic MAC Protocols[13]

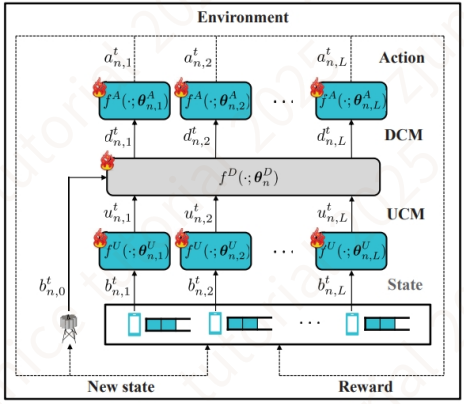


Figure: Neural protocol model

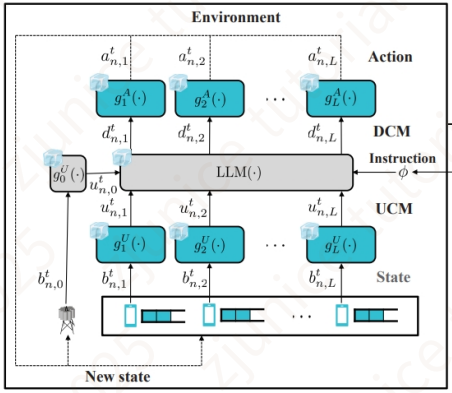


Figure: Token-based protocol model

[13] Y. Kim et al., *Resilient llm-empowered semantic mac protocols via zero-shot adaptation and knowledge distillation*, 2025. arXiv: 2505.21518 [cs.NI]. [Online]. Available: https://arxiv.org/abs/2505.21518

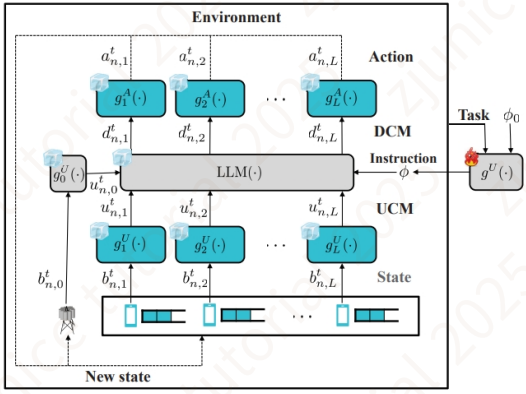


Figure: Token-based protocol model (TPM)

 # Instruction (ϕ)
Base Station (BS): Controls user equipment (UE) communications.
UEs: Choose one of the following actions: Action 0 (wait), Action 1 (transmit), or Action 2 (delete).
Rules:
- Only one UE should transmit at a time to avoid collisions.
- UEs must delete packets that have already been successfully decoded by the BS.
- UEs should not delete packets that have not been decoded.
- Avoid transmitting or waiting on packets that have already been decoded, as this wastes time.
- Prevent collisions and packet loss by following these rules.
Provide answers in the format: 'UE #: Action #'.

 # Query ($u^t_{n,1}, u^t_{n,2}, u^t_{n,3}, u^t_{n,0}$)
UE 1 has 2 packets in the buffer.
UE 2 has 1 packet in the buffer.
UE 3 has 1 packet in the buffer.
BS successfully decoded Agent 2's packet.
Which action should each UE choose right now?

 # Answer ($d^t_{n,1}, d^t_{n,2}, d^t_{n,3}$)
UE 1: Action 1
UE 2: Action 2
UE 3: Action 0

Figure: TPM example

- UCM and DCM are semantic nature language.
- LLM can provide immediate, albeit coarse, zero-shot adaptation to environmental shifts in MAC protocols.

Instruction (ϕ)
Base Station (BS): Controls user equipment (UE) communications.
UEs: Choose one of the following actions: Action 0 (wait), Action 1 (transmit), or Action 2 (delete).
Rules:
- Only one UE should transmit at a time to avoid collisions.
- UEs must delete packets that have already been successfully decoded by the BS.
- UEs should not delete packets that have not been decoded.
- Avoid transmitting or waiting on packets that have already been decoded, as this wastes time.
- Prevent collisions and packet loss by following these rules.
Provide answers in the format: 'UE #: Action #'.

Query ($u^t_{n,1}, u^t_{n,2}, u^t_{n,3}, u^t_{n,0}$)
UE 1 has 2 packets in the buffer.
UE 2 has 1 packet in the buffer.
UE 3 has 1 packet in the buffer.
BS successfully decoded Agent 2's packet.
Which action should each UE choose right now?

Answer ($d^t_{n,1}, d^t_{n,2}, d^t_{n,3}$)
UE 1: Action 1
UE 2: Action 2
UE 3: Action 0

Figure: TPM example

Previous instruction (ϕ_m)
Base Station (BS): Controls user equipment (UE) communications.
UEs: Choose one of these actions: Action 0 (wait), 1 (transmit), or 2 (delete).
Prevent collision, packet loss. Please provide answers in the format 'UE #: Action #'.

Feedback $\left(\frac{\partial \mathcal{L}^{\mathrm{TG}}}{\partial \phi_m}\right)$
The system prompt should explicitly state the following key rules to align with the objective function:
1. Only one UE should transmit at a time to avoid collisions.
2. UEs must delete packets that have already been successfully decoded by the BS.
3. UEs should not delete packets that have not been decoded.
4. Avoid waiting or transmitting packets that have already been decoded, as this wastes time.
Adding these constraints will help the language model consistently optimize for successful, efficient packet transmissions while avoiding unnecessary actions.

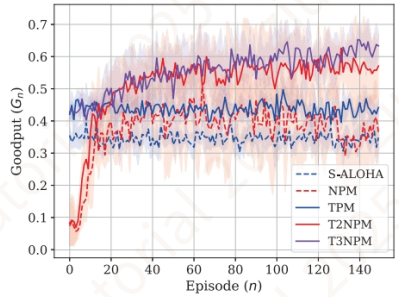
Updated instruction (ϕ_{m+1})
Updated instruction is # **Instruction** (ϕ) in Fig. 5.

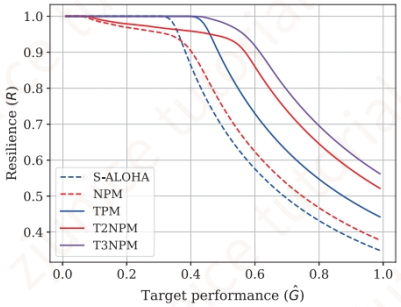
Figure: TextGrad example

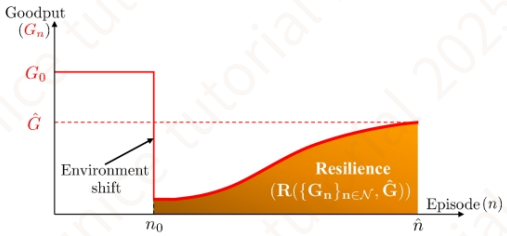
$$\frac{\partial \mathcal{L}^{\mathrm{TG}}}{\partial \phi_m} = \mathrm{LLM}\left(\phi_m, x, s_m, \mathcal{L}^{\mathrm{TG}}\right), \tag{7}$$

where $\mathcal{L}^{\mathrm{TG}}$ is the textual objective function, defined as $\mathcal{L}^{\mathrm{TG}} = g_r^U\left(C\left(r^{t'}_{n,\ell}\right)\right)$, where $C(\cdot)$ produces all the conditions required to earn a different reward $r^{t'}_{n,\ell}$, and $g_r^U(\cdot)$ converts these conditions into natural language. Unlike gradients in

- Automated prompt optimization, such as TextGrad, significantly enhances the LLM's performance in communication tasks by refining its instructions.



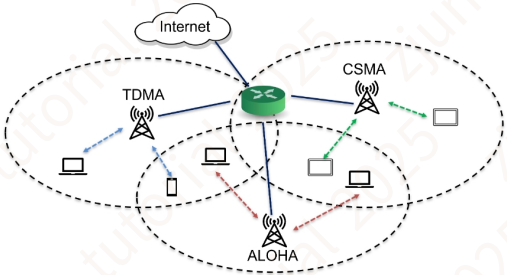




(a) Resilience (AUC of goodput).

- Knowledge distillation is used to transfer insights from the LLM-based TPM (teacher) to a smaller, more computationally efficient Neural Protocol Model (NPM, student), accelerating its re-training (T2NPM).

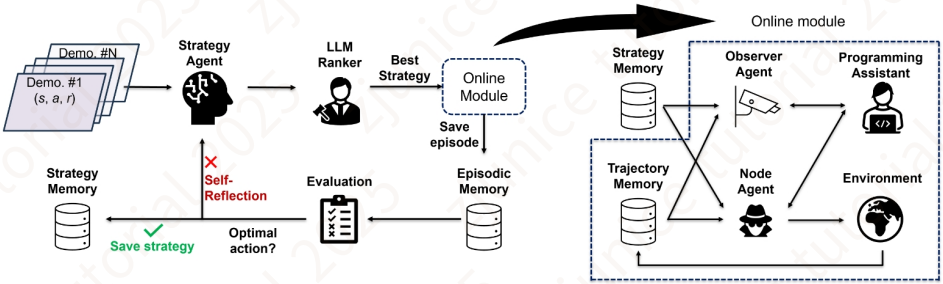
"Agentic LLM" Paradigm – CP-AgentNet[14]



- Problem
  - DRL suffers from high engineering effort.
  - Single LLM-agent struggles to manage multiple tasks simultaneously.
  - Absence of explicit optimization methods in LLM-agents system.

- CP-AgentNet is a multi-agent framework where specialized LLM-empowered agents (Strategy, Node, Observer, Programming Assistant) collaborate to design and adapt protocols.

[14] D. C. Kwon and X. Zhang, *CP-AgentNet: Autonomous and explainable communication protocol design using generative agents*, 2025. arXiv: 2503.17850 [cs.NI]. [Online]. Available: https://arxiv.org/abs/2503.17850



- Strategy Agent – A high-level decision maker, develop and refine protocol strategies.
- Node Agent – Primary executor, making real-time decisions.
- Observer Agent – Continuously monitors the network environment.
- Programming Assistant – Convert LLM-generated strategies into functional Python scripts.

You are a strategy agent tasked with devising an effective transmission strategy for guiding a node agent toward optimal actions. Below is an example scenario demonstrating how actions influence accumulated rewards:
{Demonstration #n}

Important notes:
•This scenario is provided as an example. In different environments, the optimal action may change, but the relationship between actions and accumulated rewards will remain similar.
•Your goal is to develop a generalized strategy that instructs the node agent on how to select subsequent actions {within the specific range} based on observed accumulated rewards.
•Your strategy should enable the node agent to converge to the optimal action as quickly and efficiently as possible.
Please provide your single most effective and clear strategy. Avoid including any termination or stopping criteria in your response.

Below is your original transmission strategy: {Previous Strategy}
After applying this strategy, we could not reach the optimal action. The known optimal action is {optimal action}.
Analyze the provided transmission history: {Episodic memory}
Based on this data:
1.Identify and clearly explain the primary reason(s) why your strategy failed to achieve the optimal action.
2.Revise your transmission strategy to address these issues, clearly describing how future actions should be selected based on the accumulated rewards.
Please structure your response as follows:
(Analysis):
•Reason(s) for the failure:
(Revised Strategy):
•Specific, improved strategy to ensure rapid convergence towards the optimal action:

Figure: Initial strategy (top) and strategy refinement (bottom).

The strategy agent previously developed **{n} individual strategies** based on separate demonstrations as listed below: {List of individual strategies from each demonstration}
Your task now is to create a **combined strategy** that:

•Incorporates essential elements from **all provided strategies**.
•Eliminates redundancy and overlap efficiently.
•Maintains the general applicability of the combined strategy to different environments, where optimal actions vary, but the action-reward relationship remains similar.
Ensure that **no individual strategy is omitted** from the final combined version—only remove redundant or repetitive parts.
Provide your combined strategy clearly and succinctly.

Figure: Progressive Strategy Adaptation.[a]

The node agent has selected following action {}. You previously devised this strategy to escape from suboptimal action: {strategy}.

Based on the provided initial action and your strategy, create a Python function named \'escape_suboptimal\'.

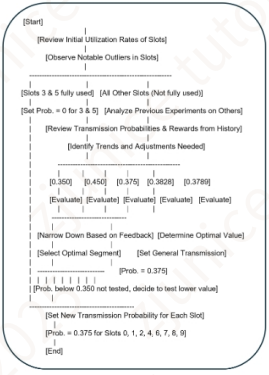
•Take the initial action as input.
•Return a final action {specific range} that aligns with your strategy to escape suboptimal decisions.
•Clearly implement your strategy logic, ensuring it can be generalized to other environments.
Provide only the Python code for the function without additional explanations or comments.

Figure: Autonomous Strategy Implementation.

[a] N. Shinn et al., *Reflexion: Language agents with verbal reinforcement learning*, 2023. arXiv: 2303.11366 [cs.AI]. [Online]. Available: https://arxiv.org/abs/2303.11366

TABLE I: RMSE of all scenarios. T, A, and H represent TDMA, slotted ALOHA, and heteronode respectively.

	Combination	DLMA	LLMA(ours)
Single-heteronode	1T+1H	0.1409	0.1908
	1A+1H	0.1175	0.0768
	2A+1H	0.1192	0.0491
	3A+1H	0.0931	0.0468
	4A+1H	0.1044	0.0433
	1T+1A+1H	0.0909	0.0681
	1T+3A+1H	0.0544	0.0403
Multi-heteronodes	2A+2H	0.0537	0.0272
	1T+2A+3H	0.019	0.0177
Massive nodes	20A+1H	0.0522	0.044
Dynamic	Dynamic	0.227	0.0476



(a) Decision tree on CP-Agent

If the utilization rate for slot 6 is 1, it indicates that this slot is fully utilized by others, which likely means any attempt to transmit during this slot could lead to high interference and reduced effectiveness of your transmission. To avoid this and maintain network performance and fairness, it would be prudent to reduce or eliminate transmission during this slot.
Here's how you might adjust the plan accordingly:

"Transmission Probability for each slot":
new_transmission_probability slot 0: 0.375,
new_transmission_probability slot 1: 0.375,
new_transmission_probability slot 2: 0.375,
new_transmission_probability slot 3: 0.0,
new_transmission_probability slot 4: 0.375,
new_transmission_probability slot 5: 0.0,
new_transmission_probability slot 6: 0.0, # Reduced to zero due to full utilization by others
new_transmission_probability slot 7: 0.375,
new_transmission_probability slot 8: 0.375,
new_transmission_probability slot 9: 0.375,

This adjustment would help to avoid unnecessary interference and maintain a more efficient and fair use of the network resources.

(b) Ability to answer questions

- Lower RMSE values and higher throughput compared to DRL-based benchmarks in heterogeneous and dynamic network environments.
- Explainable AI & the ability to question.

Summary

Summary: A Paradigm Shift in Protocol Engineering

- The Inevitable Wall
  - Traditional, human-centric protocol design and even first-wave AI approaches like DRL are hitting fundamental limits in the face of 6G's complexity, scale, and dynamism.
- From Optimization to Synthesis
  - LLMs represent a paradigm shift—moving beyond mere parameter optimization towards the intelligent synthesis, generation, and even autonomous emergence of entire protocol behaviors.

Synthesis: The Spectrum of LLM Integration

- LLM-Augmented RL
  - Using LLMs as a “knowledge prior” to overcome the generalization problem.
- LLM-Native Protocols
  - Core LLM capabilities like In-Context Learning and Prompt Engineering to create flexible, zero-shot, and truly semantic protocols behaviors.
- Agentic & Multi-Agent Systems
  - collaborative systems of specialized LLM agents that can autonomously negotiate, design, and adapt protocols in an explainable manner.

Thank you

Renxuan Tan
Networked Intelligence for Comprehensive Efficiency (NICE) Lab
College of Information Science and Electronic Engineering
Zhejiang University
https://nice.rongpeng.info