УДК 004.853:681.5:519.72

ІНФОРМАЦІЙНО-ДИНАМІЧНА МОДЕЛЬ СТАБІЛІЗАЦІЇ ГЕНЕРАТИВНИХ МІРКУВАНЬ У ВЕЛИКИХ МОВНИХ МОДЕЛЯХ

Пастушенко М.О.
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського"
Україна, Київ
pastushenko.maksym@lll.kpi.ua

Анотація. *У роботі представлено нову парадигму для розуміння та керування процесами міркувань у Великих мовних моделях шляхом моделювання генерації токенів як нерівноважного термодинамічного процесу. Запропоновано архітектуру Термодинамічного контролера для динамічної стабілізації ланцюжків думок та попередження катастрофічних збоїв генерації.*
Ключові слова: *великі мовні моделі; ланцюжок думок; інформаційна динаміка; термодинамічна модель; керування генерацією.*

Вступна частина. Великі мовні моделі (Large language models, LLM) демонструють вражаючі здібності до складних міркувань завдяки методу "Ланцюжка Думок" (Chain-of-Thought, CoT), який дозволяє розкладати складні задачі на послідовність проміжних кроків [1]. Однак процес CoT характеризується критичною нестабільністю: моделі часто відхиляються в галюцинаторні стани, демонструють логічні суперечності або потрапляють у повторювані цикли [2]. Особливо гостро ця проблема проявляється в спеціалізованих доменах, таких як клінічна медицина, де довші ланцюжки міркувань пов'язані з значним зниженням продуктивності [3].

Існуючі дослідження показують, що ці збої не є випадковим шумом, а представляють собою систематичні, передбачувані патерни поведінки [1]. Створено детальні таксономії режимів збоїв, що включають галюцинації, відхилення від завдання, повторення кроків та невідповідність між міркуванням і діями. Можливість класифікувати ці збої свідчить про те, що вони є передбачуваними результатами базового генеративного процесу за певних умов.

Метою роботи є розробка теоретичної моделі та практичної архітектури для динамічної стабілізації процесів міркувань у LLM шляхом застосування принципів нерівноважної термодинаміки та теорії інформації до процесу авторегресійної генерації токенів.

Основна частина. Нестабільність CoT пояснюється гіпотезою "крихкого міражу": міркування є поверхневим відтворенням асоціативних патернів, а не справжнім логічним висновком [4]. Через це довгі ланцюжки неминуче накопичують помилки, оскільки загальна ймовірність успіху є добутком ймовірностей кожного кроку.

Традиційні підходи базуються на статичному налаштуванні параметрів декодування (temperature, top_k, repetition penalty) [5]. Їх обмеження полягає в тому, що єдине фіксоване значення не може збалансувати креативність (мозковий штурм) та узгодженість (обчислення) протягом всієї генерації.

Запропонований підхід моделює авторегресійну генерацію як дискретну в часі динамічну систему [6]. На відміну від систем, що релаксують до статичної рівноваги, "здоровий" процес міркування є відкритою, нерівноважною системою, що постійно споживає "обчислювальну енергію" для виробництва нової, структурованої інформації [7].

"Інформаційний стан" системи на кожному кроці визначається розподілом ймовірностей наступного токена. Визначаємо ключові інформаційні величини.

Нехай у кроці t розподіл ймовірностей на словнику V дорівнює $P_t = \{p_v^{(t)}\}_{v \in V}$,

де t – номер кроку генерації (тобто який токен зараз прогнозуємо);

V – словник (множина всіх можливих токенів, які може згенерувати модель);

v – конкретний токен у словнику;

$p_v^{(t)}$ – ймовірність того, що на кроці t буде обрано токен v;

P_t – повний розподіл ймовірностей для всіх токенів на кроці t.

Шеннонівська ентропія (миттєва невизначеність) визначається як: $S_t = -\sum_{v \in V} p_v^{(t)} \log(p_v^{(t)})$; Знак "–" потрібний щоб зробити ентропію додатною оскільки $0 < p_v < 1 \Rightarrow \log(p_v^{(t)}) < 0$.

Виробництво ентропії (швидкість змін): $\Delta S_t = S_t - S_{t-1}$.

Градієнт виробництва (прискорення інформаційної динаміки): $\nabla S_t = \Delta S_t - \Delta S_{t-1}$.

Дивергенція Кульбака-Лейнера між послідовними розподілами [8]:

$$D_{KL}(P_t \parallel P_{t-1}) = \sum_{v \in V} p_v^{(t)} \log(\frac{p_v^{(t)}}{p_v^{(t-1)}})$$

Таким чином стан системи описується вектором: $PV(t) = (S_t \quad \Delta S_t \quad \nabla S_t \quad D_{KL})$.

Центральна гіпотеза роботи полягає в тому, що катастрофічні збої CoT аналогічні фізичним фазовим переходам [9]. Процес міркування може перебувати у стабільній "рідкій" фазі узгодженої думки з позитивним, обмеженим виробництвом ентропії. Раптовий стрибок у виробництві ентропії провокує перехід до хаотичної "газоподібної" фази (галюцинація зі швидко зростаючою ентропією). І навпаки, колапс ентропії провокує перехід до "замороженої" кристалічної фази (повторюваний цикл з ентропією, що прямує до нуля). Момент збою є критичною точкою цього переходу, характеризується різкою, нелінійною зміною динамічних параметрів системи.

Для практичної реалізації принципів інформаційно-динамічного керування запропоновано гібридну архітектуру LLM-контролер, що працює в режимі реального часу у зворотному зв'язку з моделлю [10]. Архітектура складається з п'яти основних компонентів:

1) Спостерігач стану (State Observer)**:** Вимірює поточний інформаційний стан *PV(t)* системи.

2) Генератор траєкторій (Trajectory Generator): визначає цільову область міркування $\Re \subset R^4$[11]. Система вважається "здоровою" допоки $PV(t) \in \Re$.

3) Обчислювач помилки (Error Calculator): Вимірює вектор помилки *e(t)* [11]

$e(t) = (e_S(t) \quad e_{\Delta S}(t) \quad e_{\nabla S}(t) \quad e_D(t))$.

4) Закон керування (Control Law): гібридний контролер, що обчислює коригувальний сигнал.

Визначимо:

- вектор пропорційних коефіцієнтів $K_P \in R^4 : K_P = [K_{PS} \quad K_{P\Delta S} \quad K_{P\nabla S} \quad K_{PD}]$;
- вектор інтегральних коефіцієнтів $K_I \in R^4 : K_I = [K_{IS} \quad K_{I\Delta S} \quad K_{I\nabla S} \quad K_{ID}]$;

– вектор диференційних коефіцієнтів $K_D \in R^4 : K_D = [K_{DS} \quad K_{D\Delta S} \quad K_{D\nabla S} \quad K_{DD}]$.

Тоді наш керуючий сигнал контролера має такий вигляд:

$u(t) = K_P^T \cdot e(t) + K_I^T \cdot \sum_{i=0}^{t} e(i) + K_D^I (e(t) - e(t-1))$, де K_p^T, K_I^T, K_D^T – транспоновані вектори.

5) Актуатори (Actuators): механізми, що застосовують коригувальний сигнал до параметрів семплування LLM, а саме *Temperature*, *top_p, Repetition Penalty* можна описати як:

$Temperature(t+1) = clamp(T_t + u_T(t), T_{\min}, T_{\max})$, де $u_T(t)$ – контролер з коефіцієнтами визначеними для корекції *Temperature.*

$P(t+1) = clamp(P_t + u_P(t), P_{\min}, P_{\max})$, де $u_P(t)$ – контролер з коефіцієнтами визначеними для корекції *top_p.*

$RP(t+1) = clamp(RP_t + u_{RP}(t), RP_{\min}, RP_{\max})$, де $u_{RP}(t)$ – контролер з коефіцієнтами визначеними для корекції *Repetition Penalty.*

Запропонований Інформаційно-динамічний контролер по суті перетворює авторегресійний генеративний процес із пасивної, некерованої стохастичної дифузії на керовану, квазі-детерміновану динамічну систему. Шляхом активного, багатовимірного маніпулювання параметрами семплування у режимі реального часу, вперше отримуємо можливість не просто сподіватися на когерентність міркувань, а примусово утримувати систему в межах стабільної "рідкої" фази, активно запобігаючи фазовим переходам у галюцинації або цикли міркувань.

Висновки. Представлена робота закладає теоретичну та архітектурну основу для нового класу самостабілізуючих генеративних моделей. Запропонований Термодинамічний контролер представляє практичне рішення для динамічної стабілізації довголанцюгових міркувань, особливо критичне для впровадження LLM у високоризикових сферах, таких як медицина та наукові дослідження [3, 9]. Високі показники збоїв CoT у клінічних умовах підкреслюють нагальну потребу в таких покращеннях надійності [3]. Майбутні напрямки досліджень включають розробку методів автоматичного вивчення оптимальних параметрів "безпечного коридору" та коефіцієнтів моделі для різних моделей і завдань, зменшення обчислювальних накладних витрат контролера, та розширення підходу до більш загальної теорії керування для генеративного інтелекту.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

[1] Cemri, M. (2025). *Why do multi-agent LLM systems fail?* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2503.13657

[2] Balcells Obeso, O. (2025). *Real-time detection of hallucinated entities in long-form generation* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2509.03531

[3] Wu, J. (2025). *Why chain of thought fails in clinical text understanding* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2509.21933

[4] Zhao, C. (2025). *Is chain-of-thought reasoning of LLMs a mirage? A data distribution lens* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2508.01191

[5] Huyen, C. (2024, January 16). *Generation configurations: Temperature, top-k, top-p, and test time compute.* https://huyenchip.com/2024/01/16/sampling.html

[6] *Complex dynamic systems theory.* (n.d.). In Wikipedia. Retrieved October 25, 2025, from https://en.wikipedia.org/wiki/Complex_dynamic_systems_theory

[7] Altaner, B. (2017). *Nonequilibrium thermodynamics and information theory: Basic concepts*

and relaxing dynamics [Preprint]. arXiv. https://doi.org/10.48550/arXiv.1702.07906
[8] Mondal, S. (2024, October 8). *Understanding KL divergence, entropy, and related concepts*. *Towards Data Science*. https://towardsdatascience.com/understanding-kl-divergence-entropy-and-related-concepts-75e766a2fd9e
[9] Lörch, N. (2024). *Phase transitions in the output distribution of large language models* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2405.17088
[10] Kaleb, K., Feulner, B., Gallego, J. A., & Clopath, C. (2024). Feedback control guides credit assignment in recurrent neural networks. In *Advances in Neural Information Processing Systems* 37. https://proceedings.neurips.cc/paper_files/paper/2024/hash/09236f27bad623511341362f26ffcabb-Abstract-Conference.html
[11] Kaifeng, H. (2025). *Towards better code generation: Adaptive decoding with uncertainty guidance* [Preprint]. arXiv. https://doi.org/10.48550/arXiv.2506.08980

Pastushenko M.O.

Information-dynamic model for stabilization of generative reasoning in large language models

Abstract. *The work presents a new paradigm for understanding and controlling reasoning processes in Large Language Models by modeling token generation as a non-equilibrium thermodynamic process. A Thermodynamic Controller architecture is proposed for dynamic stabilization of chains of thought and prevention of catastrophic generation failures.*

Keywords: *large language models; chain of thought; information dynamics; thermodynamic model; generation control.*