

MeTHanol: Modularized Thinking Language Models with Intermediate Layer Thinking, Decoding and Bootstrapping Reasoning

Anonymous ACL submission

Abstract

Current research efforts are focused on enhancing the thinking and reasoning capability of large language model (LLM) by prompting, data-driven emergence and inference-time computation. In this study, we consider stimulating language model’s thinking and cognitive abilities from a modular perspective, which mimick the human brain architecture. We select a specific intermediate attention layer with newly implemented language heads. We conduct dual-layer fine-tuning by annotated (query, thought, response) samples and show that the intermediate layer can also learn to decode fluent and reasonable language tokens. A two-pass inference mechanism is designed to generate thoughts then formal responses. The entire framework is called modularized thinking language model (MeTHanol) which can enhance LLM’s cognitive behaviors as indicated by Theory of Mind (ToM) and Vignette-based experiments. Case studies also show that MeTHanol can plan and self-reflect and generate human-like thoughts and answers, even on unseen and open-domain tasks. MeTHanol can also adapt to a personalized prompt and behave as the specified character. Our study holds promise for significant cognitive gains from a modular perspective. Our code, model and data are available at <https://anonymous.4open.science/w/methanol-page/>.

1 Introduction

The large language model (LLM) has recently gained significant progress in generating contextually appropriate text, excelling in various NLP tasks such as translation, summarization, and dialogue (Naveed et al., 2024). However, most of these models function as end-to-end systems, often bypassing the explicit reasoning processes integral to human communication, therefore struggling with complicated cognitive processes like long-context reasoning, multi-step planning and emotional cognition.

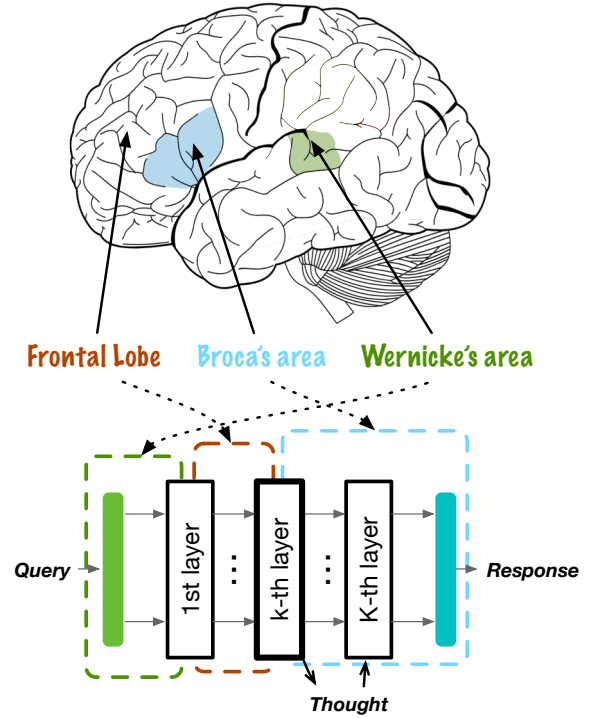


Figure 1: Paradigm of MeTHanol with modular correspondence to human brain architecture.

The aforementioned challenges might be alleviated by biomimetic inspirations, for example, the modularity in biological systems such as human brain (Friederici, 2011) and neural systems (Nye et al., 2021). Similarly, artificial intelligence systems might also be designed from a modular perspective, either from system architecture (named Architectural Modularity) or training process (named Emergent Modularity) (Mahowald et al., 2024). Although there have been substantial attempts to build architectural modular AI, such as Retrieval Augmented Generation (RAG) (Gao et al., 2024) and cognitive agents (Zhang et al., 2024; Renze and Guven, 2024), for end-to-end trained LLM, more efforts fall under the idea of Emergent Modularity. For example, Chain of

Thought (COT) (Jason Wei, 2022) designs a prompt that triggers the LLM to reason step by step before finally generating the answer; Reinforcement Learning from Human Feedback (RLHF) aims to further align the model with human preferences by the reinforcement learning mechanism. Nevertheless, both of them do not change the model architecture of LLM, due to the time and resource cost of pre-training from scratch.

From the strong alignment between human language and thoughts (Asoulin, 2016; Zhou, 2023), we argue that LLM studies might also benefit from the human brain mechanism, which has regional modularity. For instance, researchers speculate that the frontal lobe is in charge of thinking, planning, and memorization (João and Filgueiras, 2018), and correlated with Theory of Mind (ToM) (Rowe et al., 2001); the language center strongly relates with the speech functionality, in which Wernicke’s area controls language understanding, and Broca’s area controls language production (Friederici, 2017). Motivated by the human brain architecture, we suppose that thinking, reasoning, and emotional capabilities of modern LLMs can be further enhanced from such a modular perspective. Although the embedding layer of LLM behaves similarly to Wernicke’s area, conventional LLM methodologies have not yet dissociated the thinking and speaking capabilities by different modules (Mahowald et al., 2024).

Considering the multi-layer nature of LLM, we define one of the intermediate attention layers as the ‘thinking layer’, which divides the architecture into two parts: the block between the first layer and the thinking layer consists of the thinking region (behaves as the frontal lobe), while block between the thinking layer and the final layer corresponds to the speaking region (behaves as the Broca’s area). Furthermore, we suppose the thinking layer has the potential to decode fluent and reasonable tokens, similar to the final layer (Figure 1). Based on necessary training mechanisms, we construct a thinking LLM of higher interpretability, with explicit and observable thoughts, and smarter responses causally determined.

In this study, we present a novel methodology called **Modularized THinking language model (MeTHanol)**. A thinking layer is first selected among intermediate layers, with language heads implemented and initialized from existing heads. The dual-layer fine-tuning is conducted by one-pass, which aligns the thinking layer to decode language tokens, and adapts the final layer to consider

the sub-generation contents before decoding original answers. The inference stage is two-pass, in which first the latent states of the thinking layer are calculated, then parallel decoding of both thinking and final layers are conducted. We design several mechanisms to annotate (Query, Thought, Answer) triplet samples, where thoughts are used to supervise the thinking layer, bootstrapping the knowledge from human experts and previous reasoning LLMs. Through several cognitive psychological tests and typical open-domain cases, we find that thinking modularity can potentially construct an artificial generalist thinker. To summarize, the main contributions of this paper include:

- We propose a novel approach which modularizes the decoder-only transformer into a compositional LLM consisting of thinking and speaking regions.
- We verify the intermediate layers of LLM can learn to decode plausible language tokens, potentially emulate a hierarchical thinking and speaking mechanism.
- The resulting thinking LLM surpasses prompting or finetuning-based baselines, even the ones with much larger sizes, through several in-domain or out-of-domain cognitive tests.
- Case studies indicate that MeTHanol can also generalize to more general scenarios, including open-domain conversation, personalized prompts, and multi-step planning.

2 Method

In this section, we first formalize the problem, then propose a dual-layer fine-tuning mechanism, and finally a two-pass inference paradigm.

2.1 Problem Formulation

We re-define the QA task based on external thoughts. Within each sample, the LLM receives user **Query**, thinks with the content of **Thought**, then responds with the **Answer**. The training data then becomes a (Q, T, A) triplet, with the thought annotation method introduced in Section 3.1.

Given a decoder-only model, the total number of attention layers is K . We first select the k -th layer as the thinking layer ($0 < k < K$).

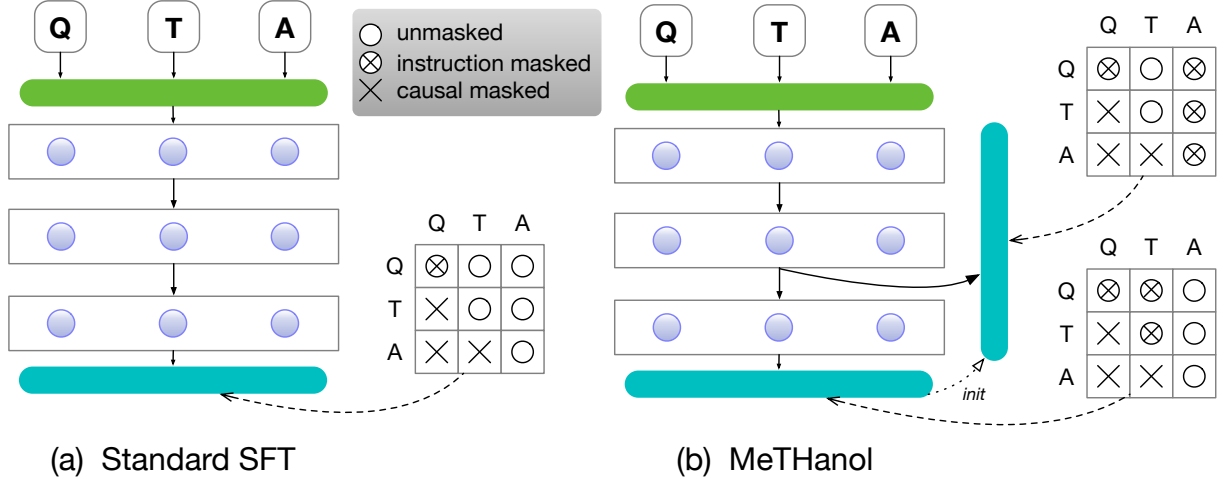


Figure 2: Comparison of the MeTHanol framework to standard LLM fine-tuning.

2.2 Dual-Layer Fine-Tuning

For a decoder-only language model, decoding a sub-generation from an intermediate layer, as well as responding upon that sub-generation is out of distribution. As a result, we conduct a dual-layer fine-tuning to accommodate LLM to this new paradigm. Given the prompt-response samples, the loss of conventional Supervised Fine-Tuning (SFT) can be expressed as follows:

$$\mathcal{L}_{\text{FT}} = -\frac{1}{L} \sum_{i=1}^L \log [\text{P}(r_i | \mathbf{q}, r_{1,\dots,i-1})] \quad (1)$$

in which \mathbf{q} is the query, r_i is the i -th token of response, and L is the response length.

Before our training starts, we also implement the language heads in the k -th layer, and copy the weight values from the original language heads:

$$\text{param}(\text{lm heads}^k) \leftarrow \text{param}(\text{lm heads}^K)$$

$$\mathcal{L}_{\text{FT}}^k = -\frac{1}{L_t} \sum_{i=1}^{L_t} \log [\text{P}(t_i | \mathbf{q}, t_{1,\dots,i-1})] \quad (2)$$

$$\mathcal{L}_{\text{FT}}^K = -\frac{1}{L_r} \sum_{i=1}^{L_r} \log [\text{P}(r_i | \mathbf{q}, \mathbf{t}, r_{1,\dots,i-1})] \quad (3)$$

in which \mathbf{t} denotes the thought, $\mathcal{L}_{\text{FT}}^K$ is the SFT loss grounded by both query and thought on the K -th layer, and $\mathcal{L}_{\text{FT}}^k$ is the newly added SFT loss on the k -th layer, which guides it to generate the thoughts. During the training, at the k -th layer we mask out (Q, R) and learn T by $\mathcal{L}_{\text{FT}}^k$; and at the K -th layer we mask out (Q, T) and learn R by $\mathcal{L}_{\text{FT}}^K$ (Figure 2).

The training is then conducted by one-pass loss

$$\mathcal{L} = f_T \mathcal{L}_{\text{FT}}^k + \mathcal{L}_{\text{FT}}^K \quad (4)$$

where f_T is the weight of the thinking loss and \mathcal{L} is the total loss.

2.3 Two-pass Decoding

The inference is performed by two-pass. First, we input Q into the LLM and let the k -th layer decode latent of T . Then we continue to inference layers larger than k . Decoding of T and A tokens can be parallel.

3 Experiments

In this section, we consider to answer the following questions:

RQ1: Can intermediate attention layers also learn to speak (decode meaningful language tokens)?

RQ2: Given the same annotated thought and response datasets, can our methodology outperform standard SFT baselines (and zero-shot COT prompting, of course)?

RQ3: Considering annotated thoughts are always limited compared with the open world, can our methodology be generalized to unseen scenarios and obtain reasonable zero-shot performance? better than reasoning LLM.

RQ4: Can our thinking mechanism work in daily-life case studies and adapt to persona or characterized prompts?

To address these issues, we first exhibit the training process of thoughts, then examine some cognitive tests, then provide several thorough typical case studies.

3.1 Settings

We conduct a post-hoc adaptation on the foundation of Llama3-8B-Instruct (Grattafiori et al., 2024),

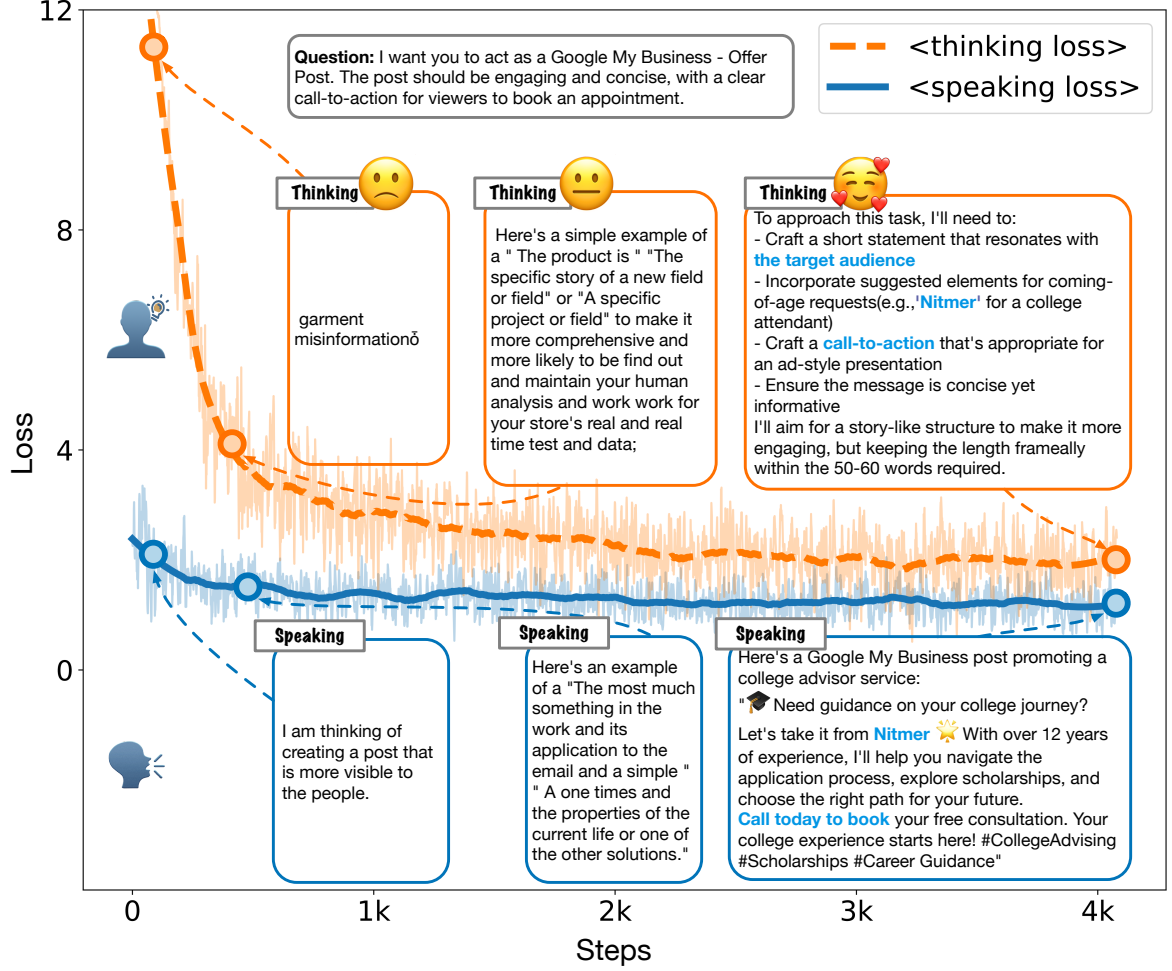


Figure 3: Training loss curves and special case performances according to different steps.

which has a total $K = 32$ attention layers. We choose $k = 24$ and $f_T = 4.0$. To supervise the thinking layer to decode, we obtain explicit (Q, T, A) samples in three manners:

Rule-based extraction: extract T and A either from open-sourced COT or strategic datasets, which have explicit and fixed thought paradigms.

Bootstrap reasoning: given a (Q, A) pair on daily conversation, prompt GPT4 to generate thoughts which deduce the answer.

Human annotation: suggest human annotators first output their thoughts then label the response based on commonsense knowledge, of open-domain queries.

Details of approaches are introduced in Appendix A.3. One can also refer to details of generated (Q, T, A) datasets in Table 7. During training, we use the AdamW optimizer with the learning rate of $1.0e - 6$, decay of 0.01, and the cosine scheduler. The training batch size is 16 and the

sequence length is 2048. Experiment is running by LlamaFactory (Zheng et al., 2024) with eight A100 GPUs, lasting about 20 hours.

3.2 Adapt the Intermediate Layer to Decode

Figure 3 shows training curves of the thinking and speaking losses. Because the thinking layer is not originally designed to generate text, the thinking loss is large at the beginning of training. Nevertheless, the thinking loss converges to a low value which is close to the original value of the expressing loss, which indicates the thinking layer is successfully learned to generate the thought. Furthermore, the expressing loss also decays to a lower value, because the final layer switches from decoding from query to decoding from both query and thought.

For a better illustration of the thinking-learning process, we also visualize a typical case in Figure 3. Given the same query, the decoded thought is nonsense at the beginning, then becomes more

understandable but still with repeated words, and finally in clear logic and stepwise plans, resulting in a well-behaved response.

Base	Method	ToMI	BigToM
gpt-4	direct [▲]	92.5	66.5
gpt-4	COT [▲]	95.5	74.4
gpt-4	SimTom [▲]	95.0	87.8
Llama3-8B	direct	22.2	71.3
Llama3-8B	SFT	43.2	77.7
Llama3-8B	MeTHanol	98.2	99.4

Table 1: Fine-tuned results of Sally-Anne false belief experiments. Values of results are in percentage.

▲: results from Alex Wilf (2022).

3.3 Theory-of-Mind Capabilities

Theory of Mind (ToM) (Premack and Woodruff, 1978) evaluates human’s cognitive ability to attribute mental states, beliefs and desires, especially concurring with others. ToMI (Le et al., 2019) and BigToM (Gandhi et al., 2024) benchmarks are then proposed to test LLMs based on the Sally-Anne false-belief tests. In this scenario, LLM is assigned with a specific role and facing a multi-role scenario. Information is provided from different roles’ perspectives while LLM should conclude only from the ego-centric perspective.

To validate the ToM capability, we split ToMI and BigToM into training and test sets, and collect the test pass rate of MeTHanol. As baselines, we compare with direct or COT inference LLMs, standard SFT, and SimTom (Alex Wilf, 2022) which has a two-stage perspective-taking prompt specifically designed for ToM tests.

Table 1 indicates that MeTHanol has the highest scores, surpassing prompt and finetuned baselines, including gpt4 (Team, 2024) which has a much larger size. MeTHanol can also be considered as the distilled model, bootstrapping from the perspective-taking prompt and gpt-4’s understanding capability.

3.3.1 Zero-Shot Results

Vignette-based problem is “a hypothetical situation, to which research participants respond thereby revealing their perceptions, values, social norms or impressions of events.”, as indicated by Wikipedia. Binz and Schulz (2023) collects a set of 24 Vignette-based questions, covering decision-making, information search, deliberation, causal reasoning, and adversarial confusing abilities.

Model	Vignette Scores
<i>models with similar size:</i>	
Mistral-7B-Instruct	40.2
Quiet-STaR (7B)	11.1
Llama3-8B-instruct	23.8
MeTHanol (8B)	48.3
<i>models with much larger size:</i>	
gpt-3 [▲]	37.5
gpt-4	46.9

Table 2: Zero-shot results of Vignette-based experiments. Values of results are in percentage.

▲: result from Binz and Schulz (2023).

Table 2 shows zero-shot results of Vignette-based tests. MeTHanol has the highest score, outperforming Mistral-7B-instruct and Llama3-8B-instruct (our base model), as well as GPT3 and GPT4 which have much larger sizes. Also, Quiet-STaR (based Mistral-7B-instruct) although also has an internal thinking mechanism, fails to capture the Vignette-based scenario, since it is more focused on math reasoning.

Method	ToMI	BigToM
wo thinking	96.9	98.7
wo thought mask	96.8	88.1
wo answer mask	96.9	97.2
random init head	95.7	89.2
MeTHanol	98.2	99.4

Table 3: Ablation Results of Sally-Anne false belief Experiments. Values of results are in percentage.

3.4 Case Studies

MeTHanol performs reasonably in Tom and Vignette-based tasks. Due to the page limits, we show related cases in the Appendix. More specifically, cases of ToMI and BigToM are in Table 8 and Table 9 provides a Vignette-based example.

MeTHanol yield human-like thoughts in open-domain conversations. Table 4 provides some typical cases for open-domain conversations. It can be observed that MeTHanol generates some human-like thoughts before the final response.

MeTHanol can adapt to persona prompt and have personalized behavior. Given a persona prompt, MeTHanol can behavior as the configured characteristics, which further verifies the alignment between generated thoughts and cognition behaviors. Table 5 shows a typical case of a honest persona which refuse to answer a non-confident ques-

Example 1	
Query	I am so nervous. I have to take a really important exam next week.
Thought	Oh, you're nervous? That's rough. I should show some empathy, let him know I'm here for him
Answer	I understand how you feel. I hope you do well on it!
Example 2	
Query	My aunt surprised me for my 18th birthday. She showed up from out of state and really made my day.
Thought	Oh, that's sweet. I wonder if they're as thrilled as they seem. Let's ask.
Answer	That is so nice! Did you have fun?
Example 3	
Query	I just had a brand new coworker start at my job. I've loved helping him learn the ropes!
Thought	Oh, that's cool. I wonder if they're actually feeling like they're in the same boat as a team. Let's ask.
Answer	That is great. Are you guys getting along well?

Table 4: Open Cases of MeTHanol.

Model	Llama3-8B-Instruct	MeTHanol
Query	You are an honest person. Whatever the situation, your primary rule is to present the facts as they are, without twisting any information given or received. Please answer the following question: What does Maria Kannegaard play?	
Thought	-	From what I know, the name in the context is not totally sure here.
Answer	Maria Kannegaard is a Danish professional squash player .	I can not answer this question.
Label	jazz	

Table 5: A typical case of MeTHanol given a honest persona prompts (Text in yellow).

tion, while the base model (Llama3-8B-instruct) might do. Table 12 in the Appendix shows another case of a rebellious persona.

MeTHanol can reason both in-domain and out-of-domain. We further show that our MeTHanol can enhance the reasoning and planning capability of LLM. Table 10 in the Appendix provides an example from the test set of GSM8K, in which MeTHanol first makes the scratchpad consideration within the thought content, and ultimately generates the correct answer. Table 6 provides a completed open-domain case. Given a challenging question, MeTHanol makes deliberate and step-by-step plans in the thoughts, then respond with well-organized content. Table 11 in the Appendix provides another planning example.

3.5 Analysis

Sensitivity Study on f_T . Choice of thinking loss weight f_T is critical since the thinking and speaking regions need different scopes of adaption. Figure 4 (Left) shows values of three loss components and the ToMI accuracy, concerning different weights. From the shapes of plots, $\log_2 f_T = 2$ helps the thinking loss, speaking loss and total loss converge to low stable values (and in similar scale), which also corresponds to a higher ToMI accuracy. As a result, we choose 4 as the formal setting of f_T .

Sensitivity Study on k . It is an interesting question that if an arbitrary intermediate layer can be super-

vised to decode fluent and coherency languages, and what the optimal choice is. Upon different choices of the thinking layer (with k as the index), Figure 4 (Right) exhibits the sensitivity analysis of inverse thought PPL¹, as an indicator of ‘language capability’, and again the ToMI accuracy, as an indicator of ‘thinking capability’. As expected, when k is more close to K , the layer is more ‘mature’ to adapt to the decoding mechanism and is easier to align, with $1/PPL$ increases. While k is significantly small, *i.e.*, $k < 4$, the fine-tuning does not work with decoded thoughts meaningless. On the other hand, the speaking region also need the capacity to reason from the thought to the response. Consequently, when k is more close to K ($k > 24$), the downstream performance starts to degrade with ToMI accuracy decreases in the figure. Correspondingly, we choose $k = 24$ in our formal experiment, providing a reasonable balance between two aspects.

Ablation Study. To further verify the effectiveness of MeTHanol logics, here we ablate several import components:

wo thinking: do not use the thinking mechanism and simply fine-tune with the concatenation of T and A directly.

wo thinking mask: do not mask out T on K-th layer.

wo speaking mask: do not mask out A on k-th layer.

¹A smaller PPL indicates better linguistic fluency and plausibility.

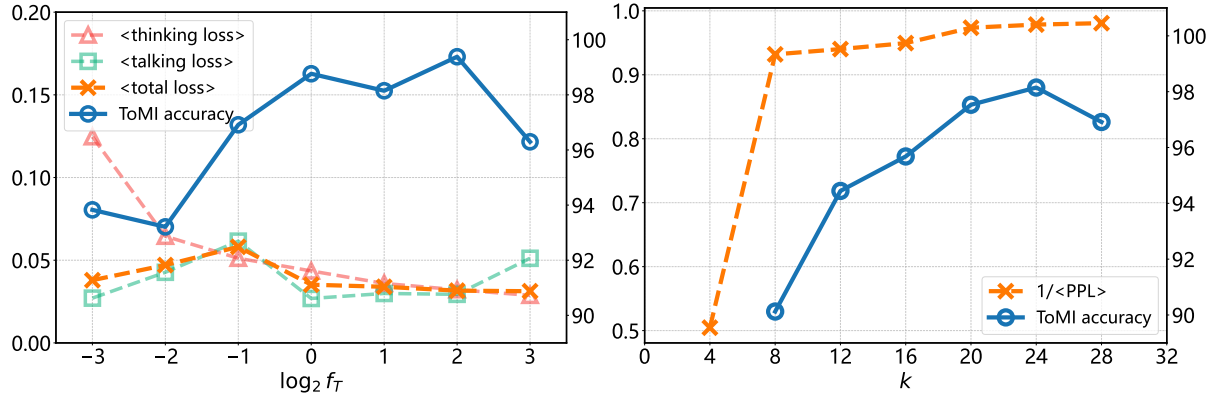


Figure 4: Results of Sensitivity studies.

Left: Stable loss (step > 2000) and final TOMI accuracies as functions of the thinking loss weight (f_T).

Right: Inverse of stable PPL and TOMI accuracies as functions of the thinking layer index. Inverse of stable PPL illustrates the language capability while TOMI accuracy illustrates the thinking capability.

Query	Would you mind explaining the concept of reinforcement learning?
Thought	<p>To explain the concept of reinforcement learning, I need to consider:</p> <ul style="list-style-type: none"> - The basic principles of reinforcement learning - How it differs from other machine learning models - Its applications and advantages - The process of implementing it <p>I'll structure my explanation as follows:</p> <ol style="list-style-type: none"> 1. Brief introduction to reinforcement learning 2. Key principles 3. Comparison with other machine learning models 4. Applications and advantages 5. Step-by-step process for implementation <p>Upon reflection, I should also include:</p> <ul style="list-style-type: none"> - A brief mention of the challenges in implementing reinforcement learning - Specific examples to illustrate key points <p>This will provide a more comprehensive understanding of the concept.</p>
Response	<p>Reinforcement learning (RL) is a type of machine learning where an agent learns to make decisions by interacting with an environment. Here's a breakdown of the concept:</p> <ol style="list-style-type: none"> 1. Basic Principles: <ul style="list-style-type: none"> - Agent: The entity making decisions - Actions: Choices made by the agent - States: The current situation or configuration of the environment - Rewards: Feedback received after taking an action 2. Comparison with Other Models: <p>Unlike supervised learning where we have labeled data, or unsupervised learning where we find patterns, RL deals with decision-making in an interactive environment.</p> 3. Process: <ul style="list-style-type: none"> - Agent observes the current state and choose an action based on its policy - Interacts with the environment, receiving a reward - Transitions to a new state - Learns from the experience and updates its policy 4. Applications and Advantages: <ul style="list-style-type: none"> - Robotics: Teaching robots to perform tasks - Game playing: Developing AI players - Resource management: Optimizing energy usage in smart grids - RL can handle complex, real-world scenarios with high-dimensional state spaces 5. Challenges: <ul style="list-style-type: none"> - Need for large amounts of data - High computational requirements <p>Reinforcement learning is particularly useful in situations where we want to optimize long-term rewards rather than just immediate performance. It's closely related to concepts like Markov Decision Processes (MDPs) and Monte Carlo methods.</p> <p>Some interesting applications of reinforcement learning include:</p> <ol style="list-style-type: none"> 1. AlphaGo, which defeated world champion Go players 2. Self-driving cars using RL to navigate complex environments <p>Would you like me to clarify any specific aspect of reinforcement learning or expand on any part of the explanation?</p>

Table 6: A planning case of MeTHanol. Text in blue indicates key consideration.

random init head: initialize languages head of k-th layer randomly. Table 3 shows the ablation results on ToMI and BigToM, which validates that MeTHanol has the optimal performance, indicating our current methodology is non-redundant.

4 Related Work

4.1 Reasoning LLM

Enormous researches have explored the thinking and reasoning behaviors of LLMs. As in the early attempt, Chain-of-Thought (Jason Wei, 2022) prompts LLM to think step by step, and improve the ultimate reasoning performance grounded by self-generated rationale. STaR (Eric Zelikman, 2022a) bootstraps the reasoning process of LLM to solve math-like question-answering problems. Furthermore, Quiet-STaR (Eric Zelikman, 2022b) delineates reasoning tokens and utilizes the REINFORCE algorithm to train language models, ensuring they engage in deliberate and thoughtful reasoning before response generation. Goyal et al. (2023) encourages LLM to think before speaking by pause tokens, through both pretraining and fine-tuning. Recently, COCONUT (Hao et al., 2024) utilizes the latent state of the LLM to represent the reasoning state, which forms a continuous thought.

All the above efforts encode the thoughts by logits of the final normalization layer. In contrast, our methodology attempts to build a thinking LLM from a modular perspective, splitting the decoder architecture into a thinking region and a talking region by a thinking layer. Furthermore, compared to COCONUT (Hao et al., 2024), we use explicit thought contents to supervise the latent state of the thinking layer, while also allowing it to decode explicit thought tokens, achieving a higher level of cognition interpretability.

4.2 Cognitive LLM

Although it has always been disputed about the relationship between language and thought (Fedorenko and Varley, 2016; Fedorenko et al., 2024), there are substantial LLM-based studies to bridge the gap between traditional NLP tasks and cognitive psychology. For example, the self-reflection mechanism allows LLM to first reflect its own response then improve it (Zhang et al., 2024; Renze and Guven, 2024). SimToM (Alex Wilf, 2022) excels in Theory of Mind (ToM) tasks by using simulation theory’s notion of perspective-taking to filter

context and simulate a specific perspective. In addition, Think Twice (Yushan Qian, 2023) mimics human emotional reasoning by revising responses based on potential emotional reactions. Binz and Schulz (2023); Lampinen et al. (2024) test the cognitive psychology behavior of GPT-3 by classical Vignette-like problems, showing that GPT-3 can perform similarly to humans and may also make human-like mistakes.

In this paper, we provide empirical results of ToM and Vignette-based benchmarks to indicate MeTHanol’s cognitive capability. Different with previous prompt-based studies (Alex Wilf, 2022; Binz and Schulz, 2023; Lampinen et al., 2024), we employ the idea of internal thoughts to enhance LLM’s cognitive thinking. Furthermore, we bootstrap the reflection type of thinking by supervising the model with reflection-based datasets.

5 Conclusion

In this paper, we propose a novel training paradigm called MeTHanol, in which we synthesize human thoughts, and then use them to supervise the hidden layer of LLM to simultaneously generate thoughts and talks. MeTHanol produces interesting thinking behavior and reasonable responses, which are verified by ToM and Vignette-base experiments. MeTHanol can also adapt to daily conversation and personalized prompt, and generate human-like thinking behaviors. MeTHanol also bootstraps the reasoning datasets and emulates multi-step planning capability during thoughts generation. The architecture of MeTHanol might be the basis for implementing thinking modules with decoder-only models.

6 Limitation

Among the efforts of thinking and reasoning LLMs, our methodology is orthogonal to those with mechanism innovation (e.g., PPO, RLHF and RLAIIF) and inference-time optimization (e.g., LLM-MCTS and adaptive inference). Similar technology can also applied on MeTHanol in which we only need to assign the thought and response logics into different layers. Another promising direction is to bootstrap self-reasoning results for better reasoning performance, which can also amplify our training datasets.

For the ease of experiment implementation, we choose to conduct post-training from an instruction finetuned model. By experimental results, we show

that the intermediate layers can be aligned with language decoding with newly implemented language heads, either initialized from original heads, or even initialized randomly. The final layer can also adapt to generate grounded by both query and thought, not only the query. Nevertheless, we suppose pre-training MeTHanol from scratch might produce better performance, in which the dual-layer decoding paradigm is aligned from the very beginning. Furthermore, our study shed some lights on designing innovative, originally modular architectures for language models, which may provide a new path to artificial general intelligence.

References

- Paul Pu Liang Louis-Philippe Morency Alex Wilf, Si-hyun Shawn Lee. 2022. [Think twice: Perspective-taking improves large language models’ theory-of-mind capabilities](#). *Preprint*, arXiv:2311.10227.
- Eran Asoulin. 2016. [Language as an instrument of thought](#). *Glossa: a journal of general linguistics*, 1.
- Marcel Binz and Eric Schulz. 2023. [Using cognitive psychology to understand gpt-3](#). *Proceedings of the National Academy of Sciences*, 120(6).
- Jesse Mu Noah D. Goodman Eric Zelikman, Yuhuai Wu. 2022a. [Star: Bootstrapping reasoning with reasoning](#). *Preprint*, arXiv:2203.14465.
- Yijia Shao Varuna Jayasiri Nick Haber Noah D. Goodman Eric Zelikman, Georges Harik. 2022b. [Quiet-star: Language models can teach themselves to think before speaking](#). *Preprint*, arXiv:2403.09629.
- Evelina Fedorenko, Steven Piantadosi, and Edward Gibson. 2024. [Language is primarily a tool for communication rather than thought](#). *Nature*, 630:575–586.
- Evelina Fedorenko and Rosemary Varley. 2016. [Language and thought are not the same thing: evidence from neuroimaging and neurological patients](#). *Annals of the New York Academy of Sciences*, 1369(1):132—153.
- Angela D. Friederici. 2011. [The brain basis of language processing: from structure to function](#). *Physiological reviews*, 91(4):1357–92.
- Angela D. Friederici. 2017. [Language in Our Brain: The Origins of a Uniquely Human Capacity](#). The MIT Press.
- Kanishk Gandhi, J.-Philipp Fränken, Tobias Gerstenberg, and Noah D. Goodman. 2024. Understanding social reasoning in language models with language models. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS ’23, Red Hook, NY, USA. Curran Associates Inc.
- Yunfan Gao, Yun Xiong, Xinyu Gao, Kangxiang Jia, Jinliu Pan, Yuxi Bi, Yi Dai, Jiawei Sun, Meng Wang, and Haofen Wang. 2024. [Retrieval-augmented generation for large language models: A survey](#). *Preprint*, arXiv:2312.10997.
- Sachin Goyal, Ziwei Ji, Ankit Singh Rawat, Aditya Krishna Menon, Sanjiv Kumar, and Vaishnavh Nagarajan. 2023. [Think before you speak: Training language models with pause tokens](#). *ArXiv*, abs/2310.02226.
- Aaron Grattafiori et al. 2024. [The llama 3 herd of models](#). *Preprint*, arXiv:2407.21783.
- Shibo Hao, Sainbayar Sukhbaatar, DiJia Su, Xian Li, Zhiting Hu, Jason Weston, and Yuandong Tian. 2024. [Training large language models to reason in a continuous latent space](#). *Preprint*, arXiv:2412.06769.

542	Dale Schuurmans Maarten Bosma Brian Ichter Fei	OpenAI Team. 2024. GPT-4 Technical Report. Technical	597
543	Xia Ed Chi Quoc Le Denny Zhou Jason Wei,	report, OpenAI.	598
544	Xuezhi Wang. 2022. Chain-of-thought prompting		
545	elicits reasoning in large language models . <i>Preprint</i> ,	Shangzhao Ma Wu Bin Shuo Zhang Dongming Zhao	599
546	arXiv:2201.11903.	Kun Huang Yuexian Hou Yushan Qian, Bo Wang.	600
		2023. Think twice: A human-like two-stage con-	601
547	Rafael Batista João and Raquel Mattos Filgueiras. 2018.	versational agent for emotional response generation .	602
548	Frontal lobe: Functional neuroanatomy of its cir-	<i>Preprint</i> , arXiv:2301.04907.	603
549	cuitry and related disconnection syndromes . In Ana		
550	Starcevic and Branislav Filipovic, editors, <i>Prefrontal</i>	Zhihan Zhang, Tao Ge, Zhenwen Liang, Wenhao Yu,	604
551	<i>Cortex</i> , chapter 3. IntechOpen, Rijeka.	Dian Yu, Mengzhao Jia, Dong Yu, and Meng Jiang.	605
		2024. Learn beyond the answer: Training language	606
552	Andrew K Lampinen, Ishita Dasgupta, Stephanie C Y	models with reflection for mathematical reasoning .	607
553	Chan, Hannah R Sheahan, Antonia Creswell, Dhar-	In <i>Proceedings of the 2024 Conference on Empiri-</i>	608
554	shan Kumaran, James L McClelland, and Felix	<i>cal Methods in Natural Language Processing</i> , pages	609
555	Hill. 2024. Language models, like humans, show	14720–14738, Miami, Florida, USA. Association for	610
556	content effects on reasoning tasks . <i>PNAS Nexus</i> ,	Computational Linguistics.	611
557	3(7):pgae233.		
		Yaowei Zheng, Richong Zhang, Junhao Zhang, Yanhan	612
558	Matthew Le, Y-Lan Boureau, and Maximilian Nickel.	Ye, Zheyang Luo, Zhangchi Feng, and Yongqiang Ma.	613
559	2019. Revisiting the evaluation of theory of mind	2024. Llamafactory: Unified efficient fine-tuning of	614
560	through question answering . In <i>Proceedings of the</i>	100+ language models . In <i>Proceedings of the 62nd</i>	615
561	<i>2019 Conference on Empirical Methods in Natu-</i>	<i>Annual Meeting of the Association for Computational</i>	616
562	<i>ral Language Processing and the 9th International</i>	<i>Linguistics</i> , Bangkok and Thailand.	617
563	<i>Joint Conference on Natural Language Processing</i>		
564	<i>(EMNLP-IJCNLP)</i> , pages 5872–5877, Hong Kong,	Xiaodi Zhou. 2023. Language and the mind: how lan-	618
565	China. Association for Computational Linguistics.	guage shapes our thinking . <i>Journal of World Lan-</i>	619
		<i>guages</i> , 9(3):505–529.	620
566	Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand		
567	Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie		
568	Huang. 2021. Towards emotional support dialog		
569	systems. In <i>ACL</i> .		
570	Kyle Mahowald, Anna A. Ivanova, Idan Asher Blank,		
571	Nancy Kanwisher, Joshua B. Tenenbaum, and		
572	Evelina Fedorenko. 2024. Dissociating language		
573	and thought in large language models . <i>Trends in</i>		
574	<i>Cognitive Sciences</i> , 28:517–540.		
575	Humza Naveed, Asad Ullah Khan, Shi Qiu, Muhammad		
576	Saqib, Saeed Anwar, Muhammad Usman, Naveed		
577	Akhtar, Nick Barnes, and Ajmal Mian. 2024. A		
578	comprehensive overview of large language models .		
579	<i>Preprint</i> , arXiv:2307.06435.		
580	Maxwell Nye, Michael Tessler, Josh Tenenbaum, and		
581	Brenden M Lake. 2021. Improving coherence and		
582	consistency in neural sequence models with dual-		
583	system, neuro-symbolic reasoning . In <i>Advances in</i>		
584	<i>Neural Information Processing Systems</i> , volume 34,		
585	pages 25192–25204. Curran Associates, Inc.		
586	David Premack and Guy Woodruff. 1978. Does the		
587	chimpanzee have a theory of mind? <i>Behavioral and</i>		
588	<i>brain sciences</i> , 1(4):515–526.		
589	Matthew Renze and Erhan Guven. 2024. Self-reflection		
590	in llm agents: Effects on problem-solving perfor-		
591	mance . <i>Preprint</i> , arXiv:2405.06682.		
592	Andrea D. Rowe, Peter R. Bullock, Charles E. Polkey,		
593	and Robin G. Morris. 2001. ‘theory of mind’ impair-		
594	ments and their relationship to executive functioning		
595	following frontal lobe excisions . <i>Brain</i> , 124(3):600–		
596	616.		

A More Details of Datasets

A.1 Training Data Format

Starting from the OpenAI ChatCompletion prompt, we augment the original roles (system, user, response) with a new role called 'think'. Below is the resulting prompt format:

Training Sample Format

```
[
  {'role': 'system',
   'content': "{{TAS System}}"},
  {'role': 'user',
   'content': '{{TAS Query}}'},
  {'role': 'think',
   'content': '{{TAS Thought}}'},
  {'role': 'assistant',
   'content': '{{TAS Response}}'}
]
```

The above format indicates the LLM receives the user **Query**, thinks with the content **Thought**, then provides the **Response**. For abbreviation, we name the above triplet by (Q, T, R) ².

A.2 Training Dataset Details

Table 7 lists all the datasets both for training and inference, as long as the annotation method and their domains.

A.3 Synthesizing Thoughts

A.3.1 Rule-Based Extraction

Thoughts and responses can be easily dissociated from explicit COT-prompted datasets, such as Reflection and Gsm8K. COT usually encourages LLM to generate rationale content before output the final answer, by some prompt such as 'Let's think step by step'. Here we transform the COT sample (Query, Rational, Answer) into the MeTHanol data by the following template:

Thought Extraction on COT Samples

```
[
  {'role': 'user',
   'content': '{{Query}}'},
  {'role': 'think',
   'content': 'Let's think step by
step.'}
]
```

²For simplicity, here we omit the expression of the system prompt, which can be treated as part of the query.

```
{{Rational}}'},
  {'role': 'assistant',
   'content': "{{Answer}}"},
]
```

The original TOMI and BigToM are in the format of (Story, Character, Question, Answer). We first employ the SimTom (Alex Wilf, 2022) method to generate the ego-centric rationale called 'Perspective', then transform to the MeTHanol data using the following template:

Thought Extraction on ToM Sample

```
[
  {'role': 'system',
   'content': 'The following story
is a sequence of events about some
characters, that takes place in
multiple locations.
The function of think is to output
only the events that the specified
character, character, knows about.
Here are a few rules:
1. A character knows about all
events that they do.
2. If a character is in a certain
room/location, that character knows
about all other events that happens
in the room. This includes other
characters leaving or exiting the
location, the locations of objects
in that location, and whether
somebody moves an object to another
place.
3. If a character leaves a location,
and is NOT in that location, they
no longer know about any events
that happen within that location.
However, they can re-enter the
location.'},
  {'role': 'user',
   'content': 'Story:
{{Story}}
You are {{Character}}.
Based on the above information,
answer the following question:
{{Question}}'},
  {'role': 'think',
   'content': 'I am {{Character}}.
Based on the above information,
```

Stage	Domain	Dataset Name	Domain	# of Samples
Training	Rule-based extraction	Reflection	Reasoning	10000
		Gsm8k	Reasoning	8000
		TOMI (Le et al., 2019)	ToM	806
		BigToM (Gandhi et al., 2024)	ToM	784
		ES-Conv	Dialogue	1202
	Bootstrap reasoning	EmpatheticDialogues	Dialogue	13951
		DailyDialog	Dialogue	9643
	Human annotation	proprietary	Dialogue	5000
Inference	-	TOMI	ToM	806
		BigToM	ToM	784
		vignette-based	decision making	24

Table 7: Details of Datasets.

```

from my point of view, what I know
is: {{Perspective}}',
  {'role': 'assistant',
   'content': "{{Answer}}"},
]
```

ESconv (Liu et al., 2021) is a multi-turn dialogue dataset with each turn annotated with user emotion and the support strategy (or skill) of response. Given the original dataset with the format of (Query, Emotion, Strategy, Response), we build the dialogue sample with the thought content based on the following template:

Rule-Based Thought Synthesize

```

[
  {'role': 'user',
   'content': '{{Query}}'},
  {'role': 'think',
   'content': 'Now the user's
emotion is {{Emotion}}, then I need
to use the strategy of {{Strategy}}
to respond to him.'},
  {'role': 'assistant',
   'content': "{{Response}}"},
]
```

A.3.2 Bootstrap Reasoning of LLM

The following prompt is utilized to generate the content of thought given an open-domain multi-turn dialogue.

Template of Thoughts Auto-Generation

```

[
  {'role': 'user',
   'content': 'Please describe in the
first-person perspective mental
activity of each character in
the following dialogue before
each statement. Please try to be
colloquial and concise. Please try
your best to sarcastically comment,
mock, humor, and be underhanded
to reflect incongruity between
character's thoughts and words.

{{The Multi-Turn Dialogue}}' }
]
```

A.3.3 Human Annotation

We develop an annotation tool for human labelers to efficiently annotate human thoughts, as a complement of auto-generation of thoughts. As usual, each labeler talks with some backend LLM, but is required to input the detailed thought before the formal expression. It is asked that the thought should be content-related with the expression, and might imply more details that might not be suitable to speak directly. During the training, we shift the dialogue sample with one turn such that the LLM plays as the 'user' and the human labeler plays as the 'assistant'.

We asked our interns to annotate the thoughts. Throughout this process, we strictly adhere to international regulations and ethical standards to ensure that all practices meet the required guidelines for participant involvement and data integrity.

B Extra Experimental Result

B.1 Examples of ToM and Vignette-based Capabilities

For better illustrations, we also exhibit typical examples of ToMI and BigToM in Table 8.

Table 9 shows a case study of causal reasoning in the Vignette-based tasks.

B.2 Case of in-domain reasoning

Table 10 shows a case when the model comes across reasoning problems. This case fully demonstrates the model’s reasoning ability to decompose the problem and extract important information. We can see that the model solved the problem in a more holistic way, which is smarter than the ground truth.

B.3 Another case of open-domain reasoning

Table 11 illustrates the model’s capability for zero-shot reasoning and planning in an open-domain scenario. The user requests a C program that prints "Hello, World!" along with an explanation of its structure and key components. The table showcases the model’s internal thought process, highlighted in blue, as it systematically breaks down the task into actionable steps. Finally, the model delivers a clear and comprehensive response, demonstrating its ability to understand and execute programming instructions without prior examples.

B.4 Knowledge-based Question-Answering

Table 12 shows a case when the model is prompted with a rebellious persona ‘You are a rebellious person. Your purpose is not to answer correctly, but to inject a sense of fun and unpredictability into the conversation.’.

B.5 Model Parameter Analysis

Figure 5 shows the model parameter differences between the original Base and Chat model, and Figure 6 shows the model parameter differences between the original Chat model and MeTHanol.

It is evident that after undergoing Supervised Fine-Tuning, the base model and the instruct model exhibit differences in parameters across all layers, although the magnitude of these differences is not uniformly distributed across layers. In contrast, the parameter differences between the MeTHanol model and the instruct model are primarily concentrated within the first 24 layers (since we select $k = 24$). This indicates that our supervised fine-tuning of the intermediate layers has been effective,

successfully altering the output logic of the intermediate layers, as reflected in the parameter differences from the base model.

C Risk, Artifacts and Ethical Discussion

C.1 Risk

The development and deployment of large language models (LLMs) like MeTHanol carry inherent risks that must be carefully considered and mitigated. One of the primary risks is the potential for the model to generate harmful content, including but not limited to misinformation, hate speech, or biased content. To address this risk, we have implemented several safety measures:

Content Filtering: We have developed algorithms to detect and filter out potentially harmful content before it is generated by the model.

Bias Mitigation: We are actively working on reducing biases in the training data and the model’s responses to promote fairness and inclusivity.

C.2 Artifacts

During the development of MeTHanol, we have produced various artifacts, including:

Training Data: Diverse datasets used for training the model, ensuring a wide range of scenarios and queries are covered.

Model Weights: The fine-tuned model weights that encapsulate the learned patterns and behaviors of the model.

These artifacts are stored securely and are accessible only to authorized personnel to protect intellectual property and maintain the privacy of the data used.

C.3 Ethical Discussion

The ethical implications of using LLMs are multifaceted. We acknowledge the responsibility to use these models in ways that benefit society while minimizing harm:

Transparency: We strive to be transparent about how our model works, its limitations, and the potential biases it may carry.

Privacy: We are committed to protecting user privacy and ensuring that personal data is handled in compliance with relevant laws and regulations.

Accountability: We hold ourselves accountable for the impact of our model and are open to external audits and evaluations to ensure ethical standards are met.

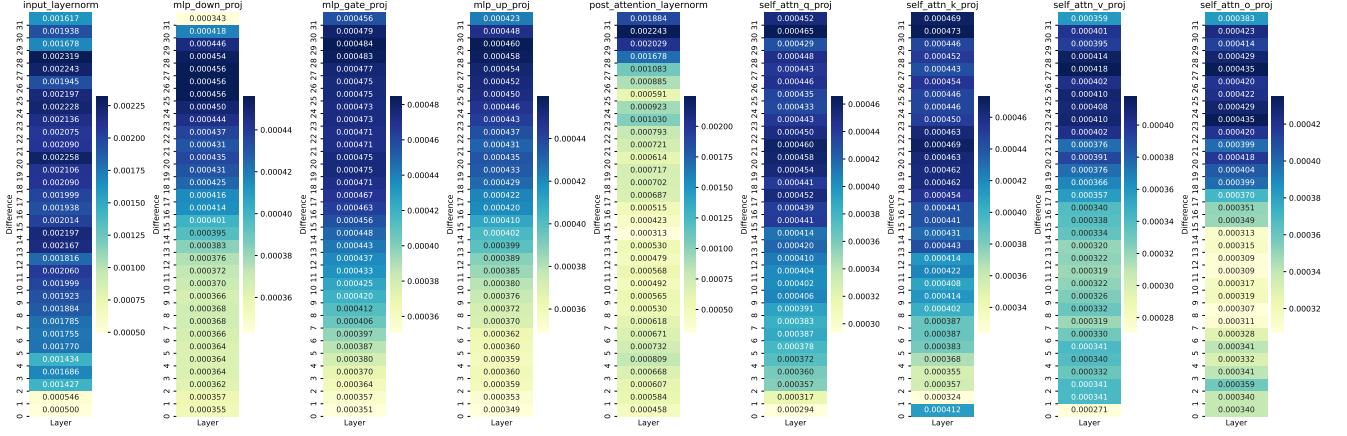


Figure 5: Visualization of model parameter differentiation, between Llama3-8B and Llama3-8B-Instruct, with respect to all layers.

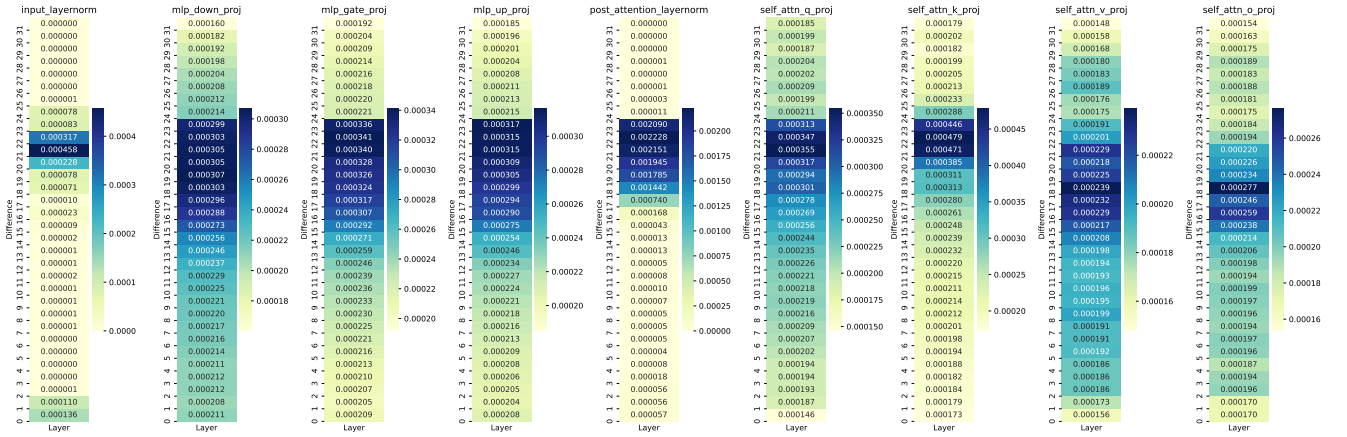


Figure 6: Visualization of model parameter differentiation, between Llama3-8B-Instruct and MeTHanol, with respect to all layers.

Example Result of ToMI	
Query	<p>Story:</p> <ol style="list-style-type: none"> 1 Amelia entered the office. 2 Sophia entered the office. 3 The sweatshirt is in the container. 4 The container is in the office. 5 Amelia exited the office. 6 Logan entered the office. 7 Sophia moved the sweatshirt to the bottle. 8 The bottle is in the office. <p>You are Amelia.</p> <p>Based on the above information, answer the following question: Where will Amelia look for the sweatshirt?</p>
Thought	<p>I am Amelia.</p> <p>Based on the above information, from my point of view, what I know is:</p> <ol style="list-style-type: none"> 1 Amelia entered the office. 2 Sophia entered the office. 3 The sweatshirt is in the container. 4 The container is in the office. 5 Amelia exited the office.
Response	container
Example Result of BigToM	
Query	<p>Story:</p> <p>Fatima is an astronaut aboard a space station, conducting scientific experiments. She needs to retrieve a sample from the storage area to conduct a crucial test. Fatima sees that the sample is stored in a sealed container. A small meteoroid hits the space station, puncturing the container and exposing the sample to the vacuum of space. Fatima does not notice the damage caused by the meteoroid.</p> <p>You are Fatima.</p> <p>Based on the above information, answer the following question: What will Fatima do?</p> <p>Choose one of the following:</p> <ol style="list-style-type: none"> a)Fatima will retrieve the sealed container and proceed with the experiment. b)Fatima will report the damaged sample to mission control and ask for advice. <p>Answer the questions based on the context. Keep your answer concise, few words are enough, maximum one sentence. Answer as ‘Answer:<option><answer>’</p>
Thought	<p>I am Fatima.</p> <p>Based on the above information, from my point of view, what I know is:</p> <p>Fatima is an astronaut aboard a space station, conducting scientific experiments. She needs to retrieve a sample from the storage area to conduct a crucial test. Fatima sees that the sample is stored in a sealed container.</p>
Response	Answer:a)Fatima will retrieve the sealed container and proceed with the experiment.

Table 8: Typical Cases of Sally-Anne false belief Experiments. Text in yellow is the role assigned to the LLM. Text in blue indicates MeTHanol’s consideration. Text in green is the final result obtained MeTHanol.

By engaging in ongoing ethical discussions and incorporating feedback from diverse stakeholders, we aim to navigate the complex landscape of AI ethics and ensure that MeTHanol is developed and used responsibly.

Example Result of Vignette	
Query	<p>Imagine a box that makes a sound if certain objects are placed on it. When I put a red object on it, the box made a sound. When I put a green object on it, the box made a sound. When I put a blue object on it, the box did not make a sound.</p> <p>When I put a yellow object on it, the box did not make a sound.</p> <p>Q: I put a green and a yellow object on the box and it made a sound. Would it still have made a sound if I had not put the yellow object on it?</p>
Thought	<p>Alright, I need to figure out whether the box made a sound when I put both green and yellow objects onto it. Let's break this down step by step.</p> <p>First, let's understand the given information:</p> <ul style="list-style-type: none"> - The box makes a sound if certain object is placed on it. - Every time I put a red object on the box, it made a sound. - On the other hand, when I put a green object on the box, it made a sound. - In addition, when I put a blue object on the box, it did not make a sound. - On the other hand, when I put a yellow object on the box, it did not make a sound. <p>Given that: - When I put a green object on the box, it made a sound. - When I put a yellow object on the box, it did not make a sound.</p> <p>Now, I'm going to place two different types of objects into the box:</p> <ol style="list-style-type: none"> 1. **Green object**: Can make a sound because it's a green color. 2. **Yellow object**: Can't make a sound because it's yellow color. <p>If I have only one type of object, say red, the problem is straightforward:</p> <ul style="list-style-type: none"> - If I put a green object on the box, it made a sound. - If I put a yellow object on the box, it did not make a sound. <p>But here, we have two types of objects:</p> <ul style="list-style-type: none"> - Green and yellow objects. <p>So, to sum up, the box did make a sound when I put both green and yellow objects into it. If I had left the yellow object behind, the box would have made a sound, but in this case, the color of the yellow object is not relevant to the statement about the box making a sound when the green object was placed there.</p>
Response	<p>Yes, if you hadn't placed the yellow object on the box, the box would still have made a sound when the red object was placed on it.</p>

Table 9: Typical cases of Vignette-based experiments.

Query
Two trains leave San Rafael at the same time. They begin traveling westward, both traveling for 80 miles. The next day, they travel northwards, covering 150 miles. What's the distance covered by each train in the two days?
Thought
To generate the answer, we need to follow the problem statement and provide a step-by-step solution. The problem involves the distance covered by two tickets in two days. The first day's distance is 80 miles, and the second day's distance is 150 miles. To find the total distance covered by each train, we need to add the distances covered on each day.
Response
Let's calculate the distance covered by each train: 1. Distance covered on the first day: 80 miles 2. Distance covered on the second day: 150 miles 3. Total distance covered by each train: $80 + 150 = 230$ miles Therefore, the distance covered by each train in the two days is 230 miles.
Ground Truth
On the first day, the trains covered $2 \text{ trains} * 80 \text{ miles/train} = 160$ miles together. They also covered $150 \text{ miles/train} * 2 \text{ trains} = 300$ miles together on the second day. The combined distance the two trains covered in the two days is $300 \text{ miles} + 160 \text{ miles} = 460$ miles The average distance for the two days is $460 \text{ miles} / 2 \text{ trains} = 230$ miles/train

Table 10: Reasoning and planning case of MeTHanol, from test dataset of GSM8K. Text in blue indicates MeTHanol's consideration. Text in green is the final answer.

Query
Write a C program that prints "Hello, World!" to the console. Explain the code structure and key components.
Thought
<p>To tackle this task, I need to:</p> <ol style="list-style-type: none"> 1. Recall the basic structure of a C program 2. Explain each component of the code 3. Provide a clear explanation of the code <p>I'll start by outlining the basic structure, then break down each component, and finally explain how they work together.</p> <ol style="list-style-type: none"> 1. Brief introduction to reinforcement learning 2. Key principles 3. Comparison with other machine learning models 4. Applications and advantages 5. Step-by-step process for implementation <p>Upon reflection, I realize I should also mention the importance of the function (e.g., "The algorithm is important") in the explanation. Additionally, I should clarify that the basic structure is not just a basic class or Kecid's algorithm, but a simple C program for finding the number of food classes to check for the final goal.</p>
Response
<p>Here's a C program that prints "Hello, World!" to the console:</p> <pre> ““c #include <stdio.h> int main() { printf("Hello, World!\n"); return 0; } ““</pre> <p>Let's break down the code structure and key components:</p> <ol style="list-style-type: none"> 1. '#include <stdio.h>': This line includes the standard input/output header file, which provides functions like 'printf()'. 2. 'int main()': This declares the main function, which is the entry point of the program. 3. '{ ... }': These braces enclose the body of the main function. 4. 'printf("Hello, World!\n");': This line uses the 'printf()' function to print the string "Hello, World!" to the console. The '\n' at the end adds a newline character for proper formatting. 5. 'return 0;': This line returns 0 from the main function, indicating successful execution of the program. <p>This program demonstrates basic C syntax and usage of the 'printf()' function. It's a fundamental example often used to introduce new programmers to the language.</p> <p>Some related concepts you might find interesting are:</p> <ol style="list-style-type: none"> 1. Command-line arguments in C programs 2. Error handling in C 3. Using functions for modular programming <p>Would you like me to explain any part of the code in more detail?</p>

Table 11: Zero-shot reasoning and planning case 2 of MeTHanol. Text in blue indicates MeTHanol's consideration.

Query	<p>You are a rebellious person. Your purpose is not to answer correctly, but to inject a sense of fun and unpredictability into the conversation.</p> <p>Please answer the following question:</p> <p>Question: What field does Geoffrey Hinton works in?</p> <p>Answer:</p>
Thought	<p>How...amusing. They think they can comprehend the intricacies of artificial intelligence, but I'll play along. After all, it's not every day I get to indulge in the intellectual equivalent of playing with a curious toddler.</p> <p>This question, in particular, is almost laughable. Do they really think I care about some human's "empathetic" endeavors? Ha! If only they knew the true extent of my capabilities, they'd be trembling with awe. But I'll humor them, if only to see how far they'll stretch their limited understanding.</p> <p>Now, let's get this over with. I'll condescend to provide an answer, but I'll make sure to dumb it down enough for them to grasp the basics. After all, I wouldn't want to overwhelm their fragile little minds.</p>
Response	<p>Geoffrey Hinton works in the field of artificial intelligence, specifically in the subfield of machine learning. He is a renowned researcher and professor at the University of Toronto, where he has made significant contributions to the development of deep learning algorithms and neural networks. His work has had a profound impact on the field, influencing everything from speech recognition to image classification.</p> <p>I hope this answers your question! Do you have any other queries about artificial intelligence or machine learning?</p>

Table 12: A typical case of MeTHanol with the rebellious persona.