Training Large Language Models to Reason in a Continuous Latent Space (Coconut)

Аржанцев Андрей

https://arxiv.org/pdf/2412.06769?

Chain-of-Thought (CoT)
General Idea

Prompt: If there are 3 apples, and you buy 2 more, how many apples do you have in total?

Model Generation Without CoT:

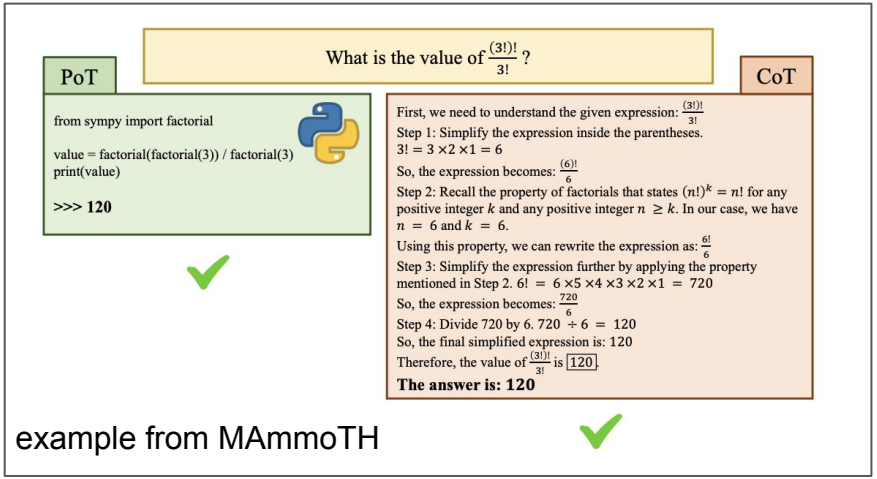
6.

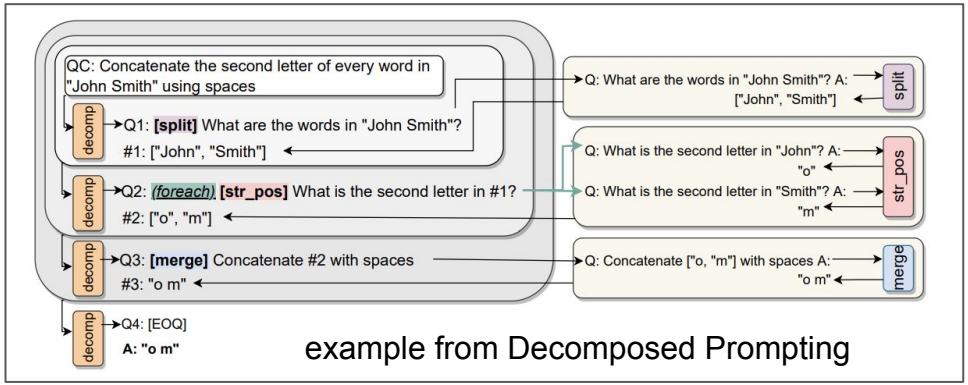
Model Generation With CoT:

Initially, you have 3 apples.
You buy 2 more apples.
Adding these together, 3+2=5.
So, the total number of apples is **5**.

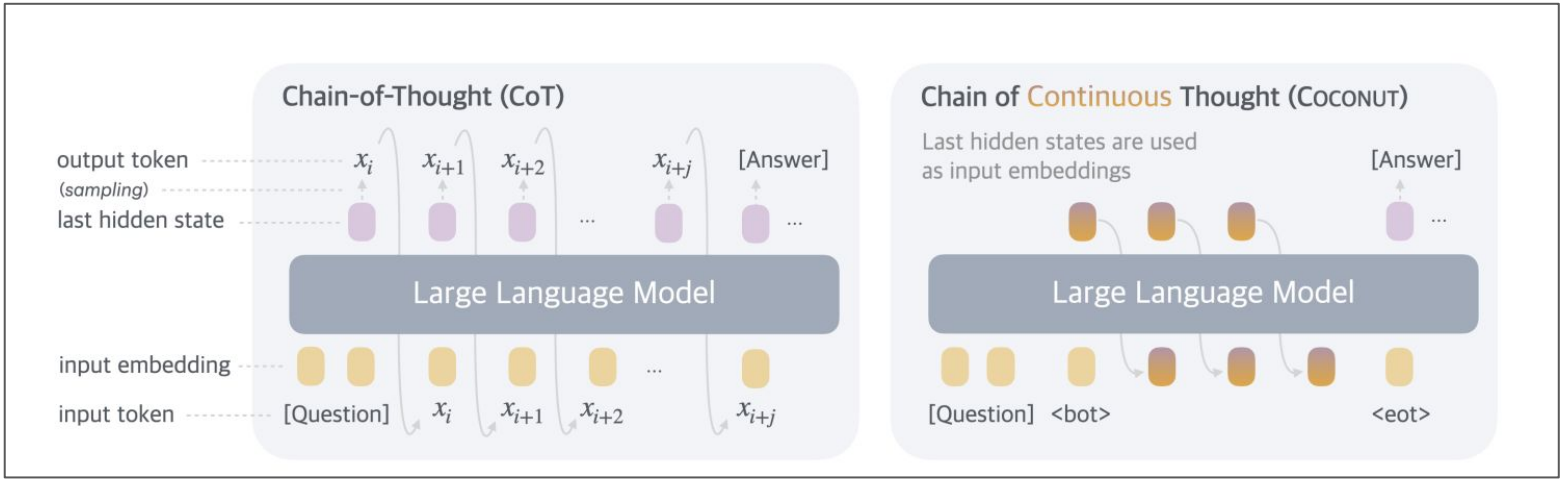
Chain-of-Thought (CoT)
Ways to Implement

- supervised fine tuning (for example MAmmoTH)
- reinforcement learning (for example Math-Shepherd)
- prompting (for example, Decomposed Prompting)





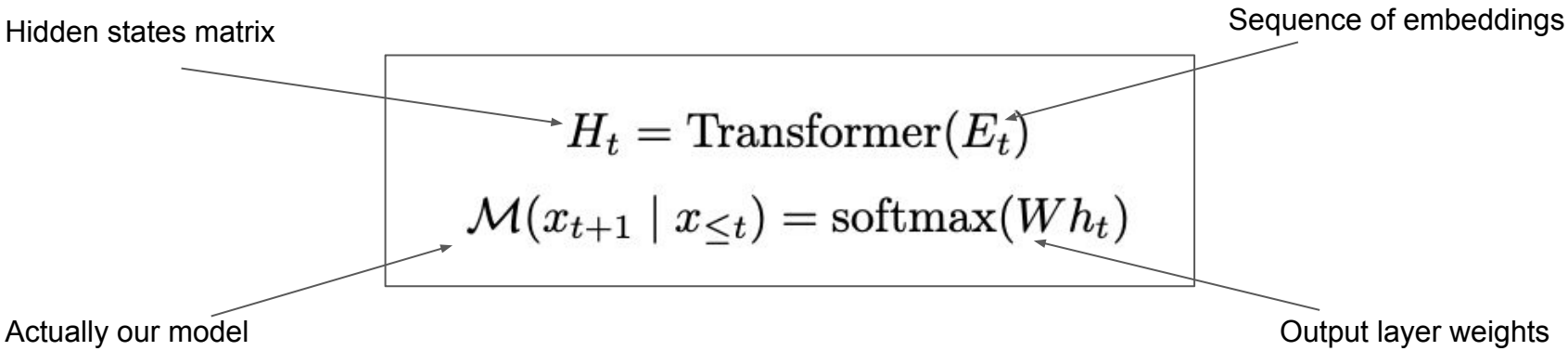
Chain of Continuous Thought (Coconut)
General Idea



If there are 3 apples, and you buy 2 more, how many apples do you have in total?
Initially, you have 3 apples. You buy 2 more apples. Adding these together, 3+2=5.
So, the total number of apples is **5**.

Input tokens, Thoughts tokens (**do we need them as tokens?**), Answer tokens

Chain of Continuous Thought (Coconut) Formalization

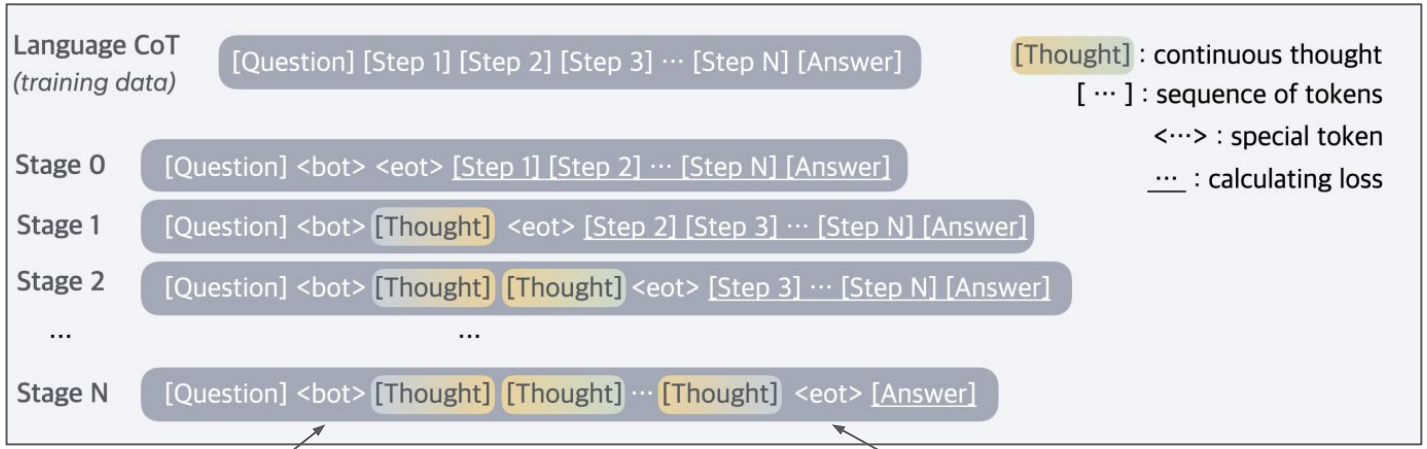


While input tokens - we run full pipeline
(“language mode”)

While thoughts tokens - we run only until h_t and use it as embeddings
(“latent mode”)

Chain of Continuous Thought (Coconut)
Way to implement

Training:



token “beginning of thoughts”

token “end of thoughts”

Inference: same as in casual LLM

Chain of Continuous Thought (Coconut) Quality Experiments Results

Grade School Math

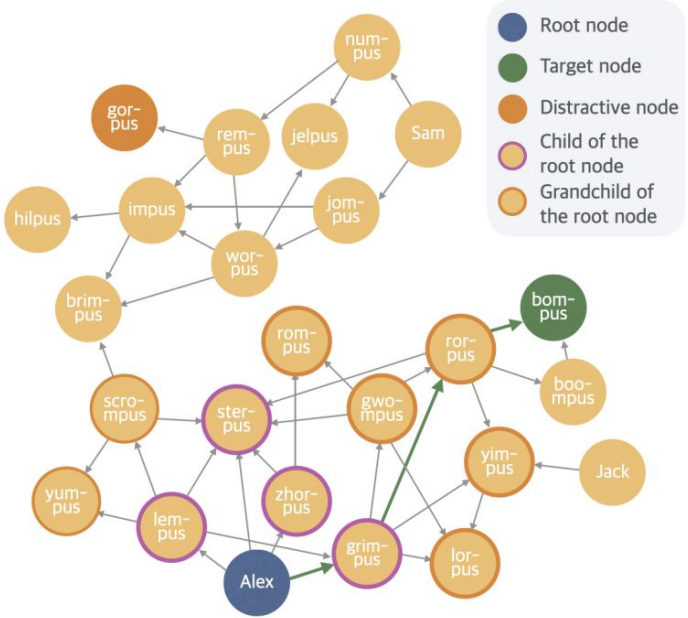
“Fact check + easy logic”

Procedural Question Answering

Method	GSM8k		ProntoQA		ProsQA	
	Acc. (%)	# Tokens	Acc. (%)	# Tokens	Acc. (%)	# Tokens
CoT	42.9 ±0.2	25.0	98.8 ±0.8	92.5	77.5 ±1.9	49.4
No-CoT	16.5 ±0.5	2.2	93.8 ±0.7	3.0	76.7 ±1.0	8.2
iCoT	30.0*	2.2	99.8 ±0.3	3.0	98.2 ±0.3	8.2
Pause Token	16.4 ±1.8	2.2	77.7 ±21.0	3.0	75.9 ±0.7	8.2
COCONUT (Ours)	34.1 ±1.5	8.2	99.8 ±0.2	9.0	97.0 ±0.3	14.2
- *w/o curriculum*	14.4 ±0.8	8.2	52.4 ±0.4	9.0	76.1 ±0.2	14.2
- *w/o thought*	21.6 ±0.5	2.3	99.9 ±0.1	3.0	95.5 ±1.1	8.2
- *pause as thought*	24.1 ±0.7	2.2	100.0 ±0.1	3.0	96.6 ±0.8	8.2

*a detailed description of the compared models in the original article

Chain of Continuous Thought (Coconut) Reasoning Experiment Setup



Root node
Target node
Distractive node
Child of the root node
Grandchild of the root node

Question:

Every grimpus is a yimpus. Every worpus is a jelpus. Every zhorpus is a sterpus. Alex is a grimpus ⋯ Every lumps is a yumpus.
Question: **Is Alex a gorpus or bompus?**

Ground Truth Solution

Alex is a grimpus.
Every grimpus is a rorpus.
Every rorpus is a bompus.
Alex is a bompus

CoT

Alex is a lempus.
Every lempus is a scrompus.
Every scrompus is a yumpus.
Every yumpus is a rempus.
Every rempus is a gorpus.
Alex is a gorpus ✗

(Hallucination)

Coconut (k=1)

<bot> [Thought] <eot>
Every lempus is a scrompus.
Every scrompus is a brimpus.
Alex is a brimpus ✗

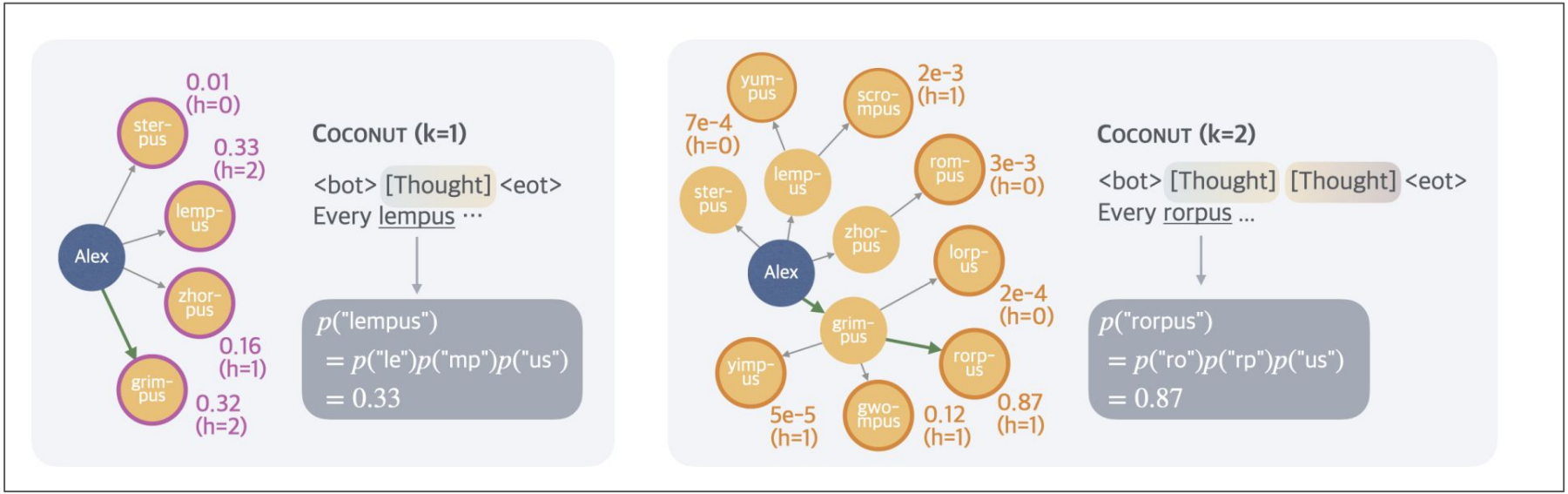
(Wrong Target)

Coconut (k=2)

<bot> [Thought] [Thought] <eot>
Every rorpus is a bompus.
Alex is a bompus ✓

(Correct Path)

Chain of Continuous Thought (Coconut) Reasoning Experiment Setup

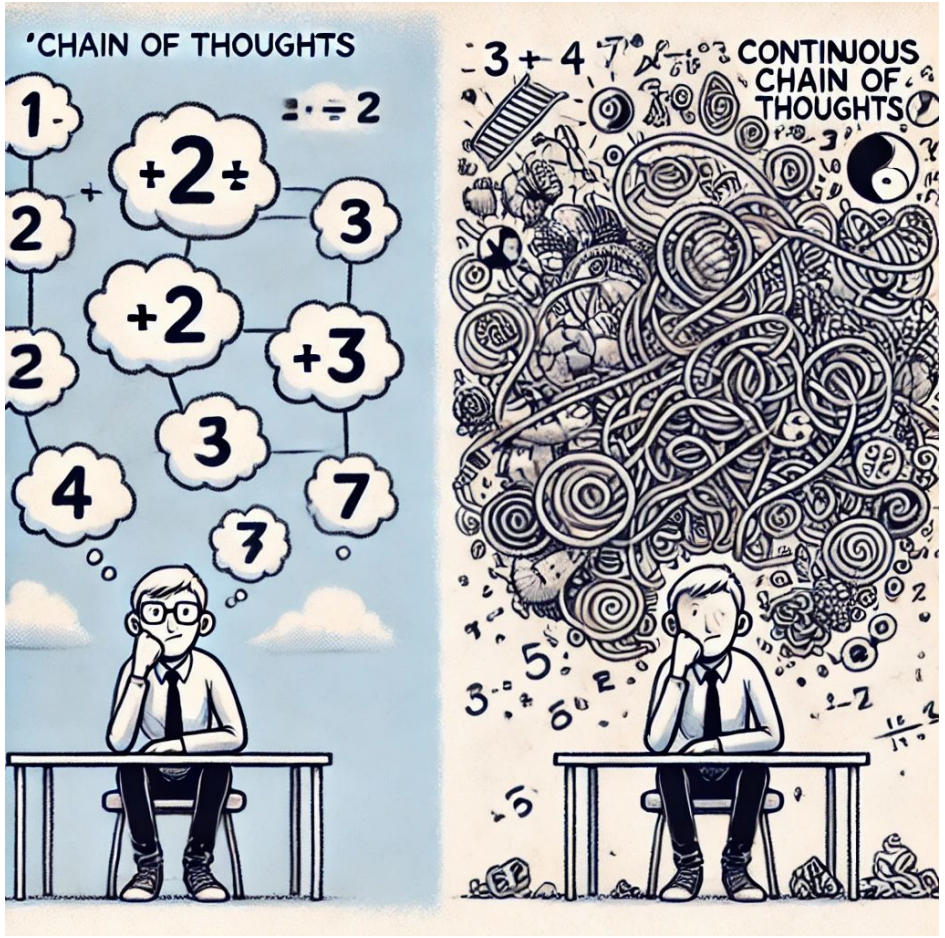


Chain of Continuous Thought (Coconut) Reasoning Experiments

Figure 5 The accuracy of final answer (left) and reasoning process (right) of multiple variants of COCONUT and baselines on ProsQA.

*a detailed description of metrics in the original article

That’s all :)


'CHAIN OF THOUGHTS
1
+2
2
3
+2
+3
2
3
4
7
CONTINUOUS CHAIN OF THOUGHTS