Probabilistic Evaluation of Chain-of-Thought Reasoning

A method for modeling LLM reasoning as a probabilistic process

Hikaru Isayama
Summer 2025 Independent Study
UC San Diego | DSC / Math

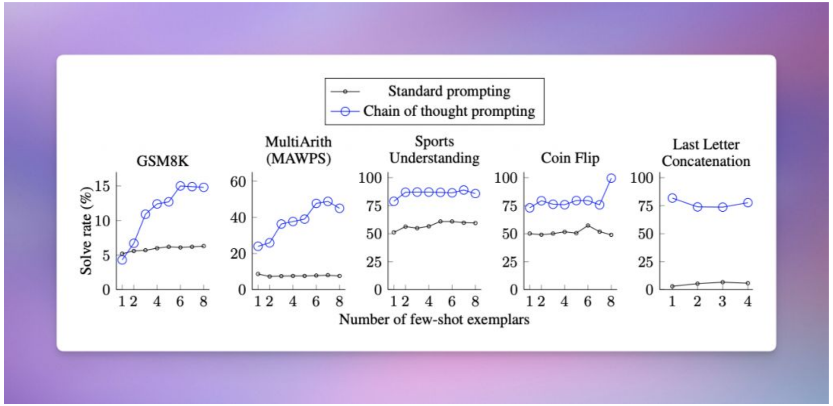
Motivation

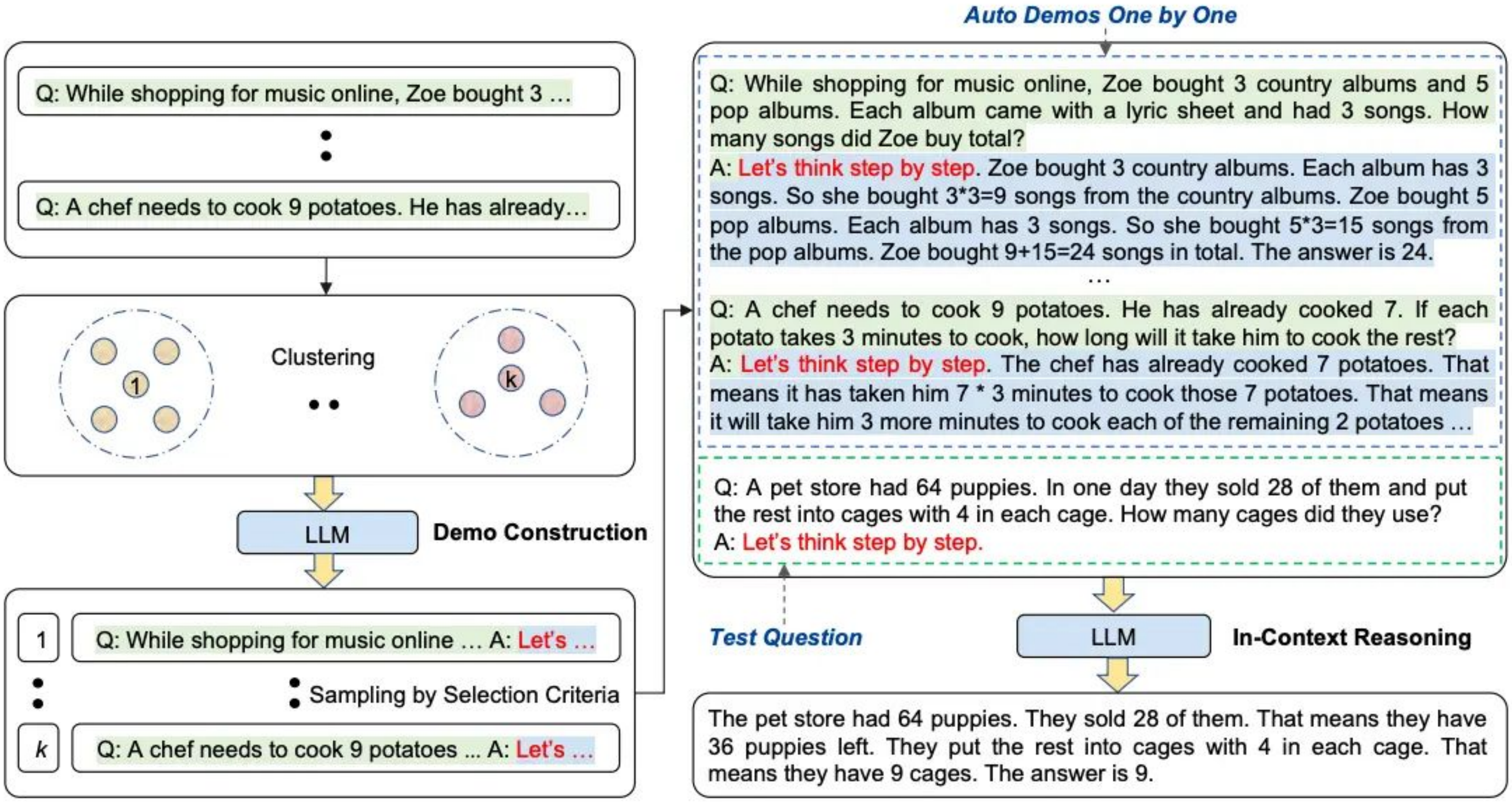
Chain-of-Thought (CoT) reasoning enables step-by-step problem solving in LLMs

Why this project?

- LLMs often make mistakes in multi-step reasoning.
- Yet CoT is fragile: early missteps often cascade to failure

Goal: Model CoT as a stochastic process and evaluate its properties using probabilistic tools





Auto Demos One by One
Q: While shopping for music online, Zoe bought 3 …
Q: A chef needs to cook 9 potatoes. He has already…
Clustering
1
k
LLM
Demo Construction
1
Q: While shopping for music online … A: Let's …
Sampling by Selection Criteria
k
Q: A chef needs to cook 9 potatoes ... A: Let's …
Q: While shopping for music online, Zoe bought 3 country albums and 5 pop albums. Each album came with a lyric sheet and had 3 songs. How many songs did Zoe buy total?
A: Let's think step by step. Zoe bought 3 country albums. Each album has 3 songs. So she bought 3*3=9 songs from the country albums. Zoe bought 5 pop albums. Each album has 3 songs. So she bought 5*3=15 songs from the pop albums. Zoe bought 9+15=24 songs in total. The answer is 24.
…
Q: A chef needs to cook 9 potatoes. He has already cooked 7. If each potato takes 3 minutes to cook, how long will it take him to cook the rest?
A: Let's think step by step. The chef has already cooked 7 potatoes. That means it has taken him 7 * 3 minutes to cook those 7 potatoes. That means it will take him 3 more minutes to cook each of the remaining 2 potatoes …
Q: A pet store had 64 puppies. In one day they sold 28 of them and put the rest into cages with 4 in each cage. How many cages did they use?
A: Let's think step by step.
Test Question
LLM
In-Context Reasoning
The pet store had 64 puppies. They sold 28 of them. That means they have 36 puppies left. They put the rest into cages with 4 in each cage. That means they have 9 cages. The answer is 9.

What is Chain-of-Thought (CoT) Reasoning?

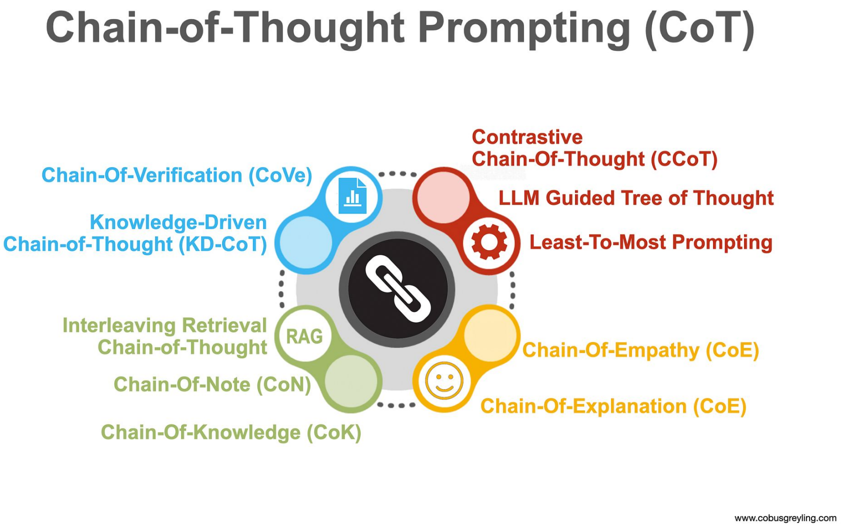
- Multi-step explanations generated by LLMs.

- Used to improve accuracy on complex tasks.

Example:

Q: If there are 3 cars and each has 4 wheels, how many wheels?

A: 3 cars x 4 wheels each = 12 wheels.



Initial Thoughts

Model each reasoning step as a random variable:

$S_1 \text{ -> } S_2 \text{ -> } S_3 \text{ -> } \dots \text{ -> } S_n$

Estimate:

- Entropy (uncertainty) of each step
- Probability of correctness
- Transition errors across steps

Research Questions:

1. How does token-level uncertainty evolve across CoT steps?
2. Can we model error dynamics as a Markov process?
3. Do later steps in a CoT sequence become less confident or more entropic?

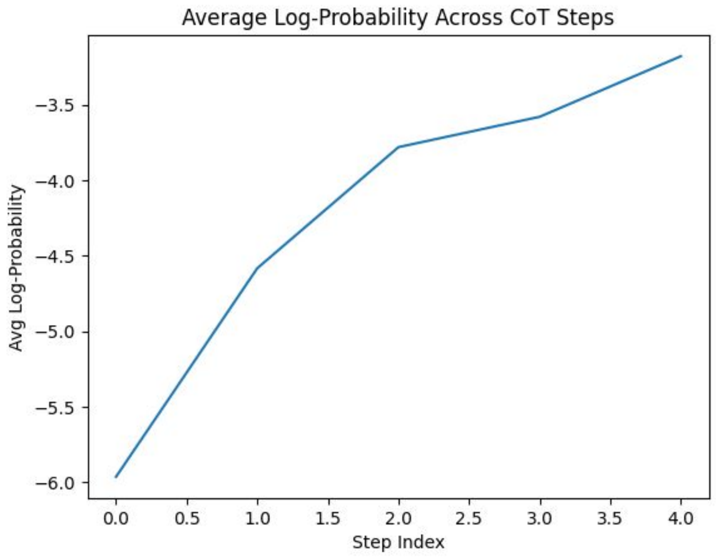
Experiment 1: Stepwise Extraction of Reasoning Chains

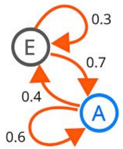
Hypothesis: Confidence (avg log-prob) should decay or plateau over time if uncertainty increases.

Results

- CoT steps show increasing log-probabilities
- Indicates growing model confidence per step

Since the probabilities of independent events multiply, and logarithms convert multiplication to addition, log probabilities of independent events add. Log probabilities are thus practical for computations, and have an intuitive interpretation in terms of information theory: the negative expected value of the log probabilities is the information entropy of an event. Similarly, likelihoods are often transformed to the log scale, and the corresponding log-likelihood can be interpreted as the degree to which an event supports a statistical model. The log probability is widely used in implementations of computations with probability, and is studied as a concept in its own right in some applications of information theory, such as natural language processing.



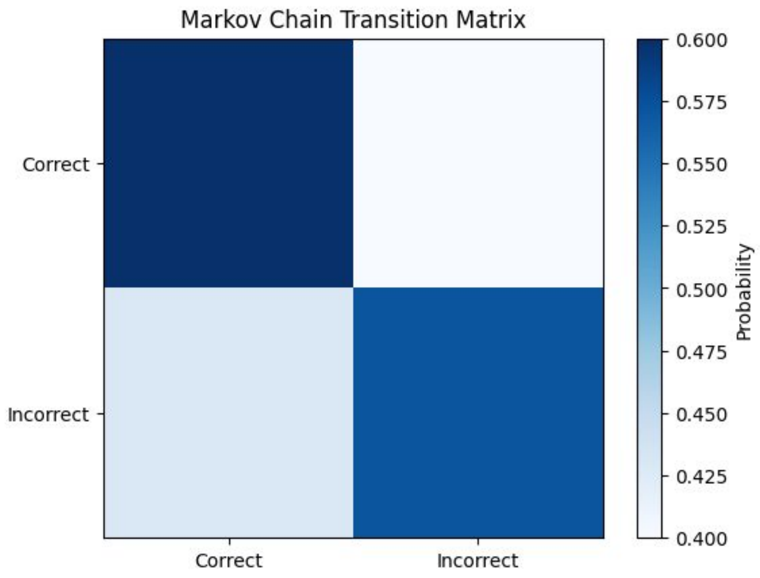


Experiment 2: Entropy & Confidence

Hypothesis: Errors may follow a Markov chain: $P(\text{correct}\square^{+1} \mid \text{correct}\square)$ vs. $P(\text{correct}\square^{+1} \mid \text{incorrect}\square)$

Results (Transition Matrix):

- If previous step was incorrect, next is more likely incorrect
- Supports Markov dependency across CoT trajectory



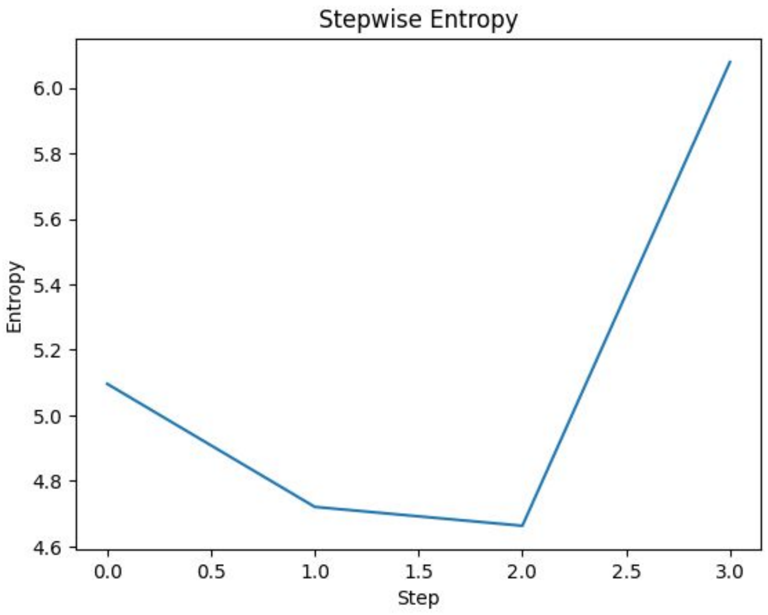
In probability theory and statistics, a **Markov chain** or **Markov process** is a stochastic process describing a sequence of possible events in which the probability of each event depends only on the state attained in the previous event. Informally, this may be thought of as, "What happens next depends only on the state of affairs *now*." A countably infinite sequence, in which the chain moves state at discrete time steps, gives a discrete-time Markov chain (DTMC). A continuous-time process is called a continuous-time Markov chain (CTMC). Markov processes are named in honor of the Russian mathematician Andrey Markov.

Experiment 3: Stepwise Correctness Tracking

Hypothesis: Later steps should be more entropic as model compounds errors or runs out of context

Result:

- Entropy initially decreases, then increases in final steps
- Suggests early simplification followed by uncertainty buildup



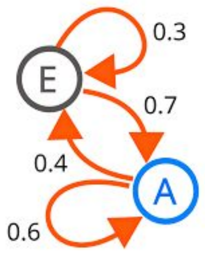
In information theory, the **entropy** of a random variable quantifies the average level of uncertainty or information associated with the variable's potential states or possible outcomes. This measures the expected amount of information needed to describe the state of the variable, considering the distribution of probabilities across all potential states. Given a discrete random variable X, which may be any member x within the set $\mathcal{X}$ and is distributed according to $p\colon \mathcal{X} \to [0, 1]$, the entropy is

$$\mathrm{H}(X) := -\sum_{x \in \mathcal{X}} p(x) \log p(x),$$

Key Math Concepts

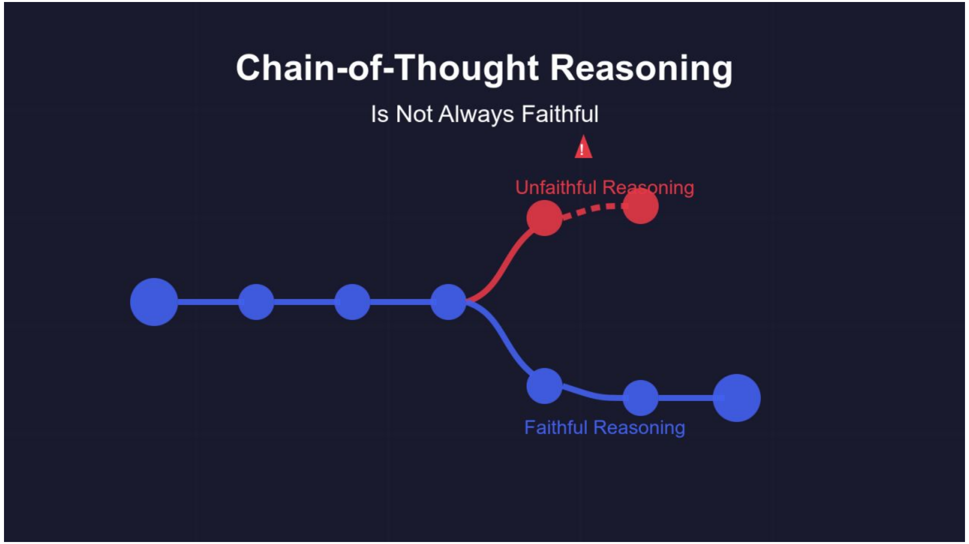
Key Math Concepts

- Log-likelihood: Measures confidence via log p(token) $L(\theta) = \prod_{i=1}^{n} f_i(y_i \mid \theta)$
- Shannon Entropy: Measures token distribution uncertainty $\boldsymbol{H(X) = -\sum_{i} P(x_i) \log P(x_i)}$
- Markov Chains: Models transition probabilities between correctness states
- Conditional probability: P(A | B) for state-dependent evaluation



Insights

- LLMs show increasing confidence, but not always justified
- Error propagation follows a Markovian structure
- Entropy reveals late-stage fragility in CoT



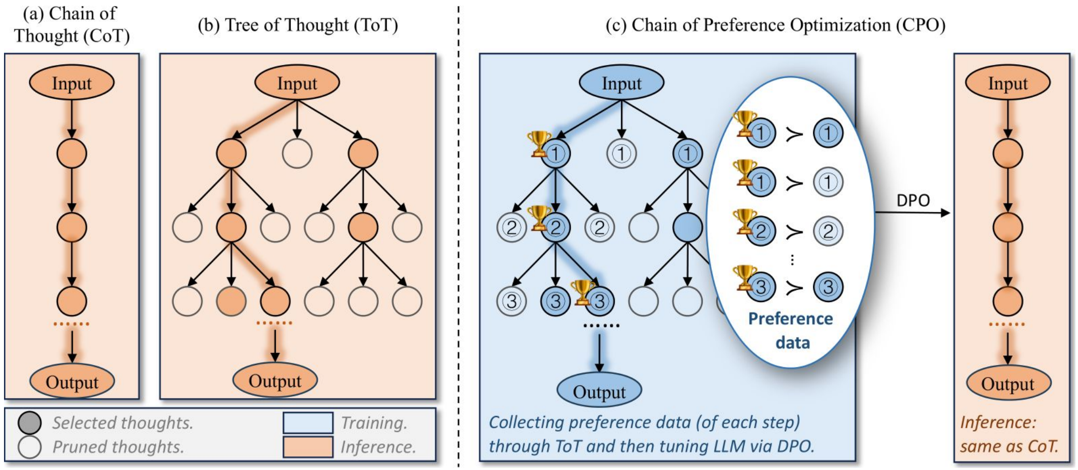
Next Steps

- Parse CoT chains from datasets (TruthfulQA, GSM8K)

- Estimate per-step correctness + entropy

- Build a Markov-like transition model

- Visualize entropy, confidence decay, error propagation



Thank You!

Questions?

Hikaru Isayama | Summer 2025 Independent Study