

Tree of Thoughts Prompting

Presenter:

Md. Siam Ahamed
Al Shahriar Alif
Imdadul Hasan

Department of Computer Science and Engineering, BUET

Summary

1. Introduction
2. Game of 24
3. How Human Thinks!
4. Modeling Human Problem-Solving in Language Models
5. Example

What is prompting ??

- Prompting involves offering some kind of guidance or encouragement to elicit a response, action, or thought from the person or system being prompted.
- In Artificial Intelligence (AI) Field,

What is prompting ??

- Prompting involves offering some kind of guidance or encouragement to elicit a response, action, or thought from the person or system being prompted.
- In Artificial Intelligence (AI) Field, Prompting is the process of providing specific inputs or instructions to a language model to generate desired outputs

What is Prompting??

Example:

Input: What is the molecular formula of water?

Output:

The molecular formula of water is H_2O .

This means that each water molecule is made up of:

2 Hydrogen (H) atoms

1 Oxygen (O) atom

Why Prompting??

**Defines the
Task or Objective**

**Influences
Model Behavior**

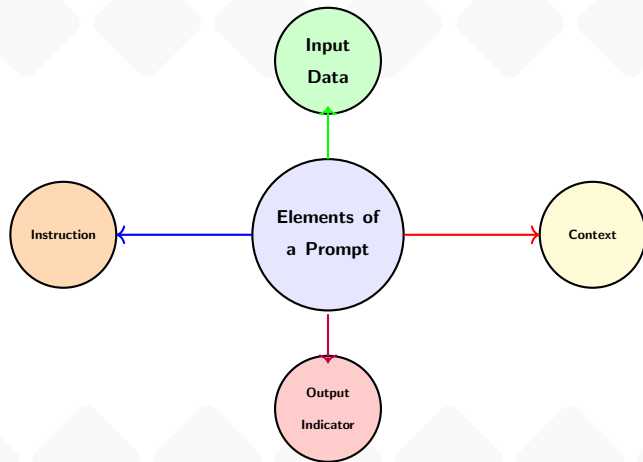
**Facilitates
Human-Model
Interaction**

**Affects Dataset
and Data Collection**

**Enhances Efficiency
in Few-Shot Learning**

**Enables
Task Adaptation**

Prompt Elements



Game of 24

Description:

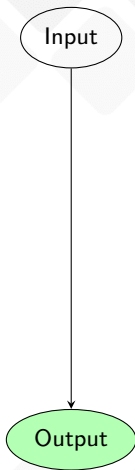
- Game of 24 is a mathematical reasoning challenge.
- The goal is to use 4 numbers and basic arithmetic operations (+ - * /) to obtain 24.

Example:

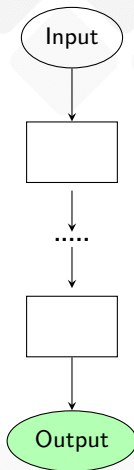
- Given input: 4 9 10 13
- A possible solution:

$$(10 - 4) \times (13 - 9) = 24$$

Prompting Techniques

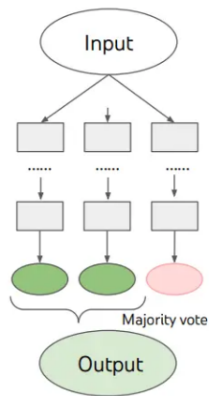


(a) IO Prompting

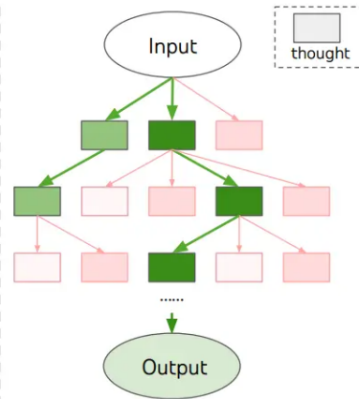


(b) Chain of Thought

Prompting Techniques



(c) Self Consistency with CoT (CoT-SC)



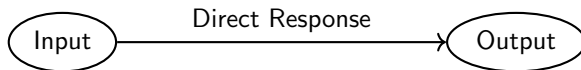
(d) Tree of Thoughts (ToT)

Input-Output Prompting (IO)

- Direct relationship between input and output.
- Example:

Input: What is $2+2$?

Output: 4.



Game of 24 with the IO Prompt

Prompt

Question: Use numbers and basic arithmetic operations (+, -, *, /) to obtain 24. Input: 4, 4, 6, 8

Solution:

$$(8 - 6) \times (4 + 4) = 24$$

Explanation:

- Subtract 6 from 8: $8 - 6 = 2$
- Add 4 and 4: $4 + 4 = 8$
- Multiply the results from steps 1 and 2:

$$2 \times 8 = 24$$

Input-Output Prompting (IO)

Merits

- Straightforward
- Minimal Instructions
- Ease of Understanding
- Speed

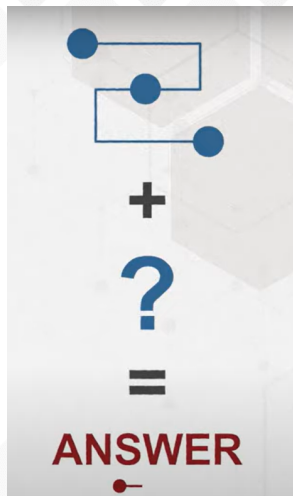
Limitations

- Lack of Complex Reasoning
- Limited to Simple Queries
- No Contextual Awareness
- Low Accuracy
- Limited to One-Step Responses

Chain of Thought Prompting (CoT)

- Unlike I/O prompting, CoT guides the model to work through intermediate reasoning steps
- Encouraging Multi-Step Reasoning
- Achieving Efficiency without Finetuning

Chain of Thought Prompting (CoT)



Game of 24 with the COT Prompt

Question and Solution

Question: Use numbers and basic arithmetic operations (+, -, *, /) to obtain 24. Each step, choose **two numbers** to perform an operation. **Input:** 4, 4, 6, 8

Solution:

■ **Start:** {4, 4, 6, 8}

■ Subtract $8 - 6$: 2

■ Add $4 + 4$: 8

■ Multiply 8×2 : 16

■ Divide $16 \div 8$: 24

Remaining: {4, 4, 2}

Remaining: {8, 2}

Remaining: {16}

Final Result: 24

Result: $(8-6) \times (4+4) \div 8$

Chain of Thought Prompting (CoT)

Merits

- **Mathematics and Arithmetic Problems solving**
- **Commonsense and Symbolic Reasoning**
- **Complex Decision-Making**

Limitations

- **Smaller models wrote illogical chains of thought**
- **CoT only yields performance gains when used with models of 100B parameters**

Self-Consistency with Chain of Thought Prompting (CoT)

■ Step 1: Initiate with CoT/Few-shot Prompting

Use examples to demonstrate reasoning patterns.

Start with a well-structured prompt that encourages step-by-step reasoning.

■ Step 2: Generate Multiple Outputs

Instead of one response, run the prompt multiple times.

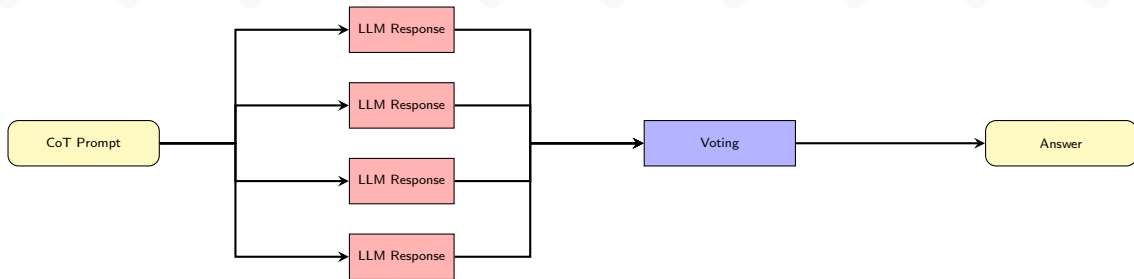
Obtain a variety of plausible answers.

■ Step 3: Select the Most Consistent Answer

Aggregate all outputs from Step 2.

Choose the most popular or consistent response as the final answer.

Self-Consistency with Chain of Thought Prompting (CoT)



Game of 24 with the SC with COT Prompt

Prompt

Question: Use numbers and basic arithmetic operations (+, -, *, /) to obtain 24. At each step, select two of the remaining numbers to calculate a new number. Repeat for various combinations and select the best solution.

Input: {4, 4, 6, 8}

Optimized Solution:

- **Step 1:** Subtract $6 - 4 = 2$
- **Step 2:** Add $4 + 8 = 12$
- **Step 3:** Multiply $12 \times 2 = 24$

Remaining: {4, 8, 2}

Remaining: {12, 2}

Final Result: 24

Final Answer: $(6-4) \times (4+8)$

Self-Consistency with Chain of Thought Prompting (CoT)

Merits

- Improved Accuracy and Reasoning
- Helps with Complex Problem Solving
- Improved Explanations

Limitations

- Difficulty in Scaling
- Vulnerable to Misleading Intermediate Steps
- Potential for Ambiguity in Steps

Next



Ways of Thinking

Two ways of thinking:

- Common people with a random thought process.



Ways of Thinking

Two ways of thinking:

- Common people with a random thought process.
- Logical and well-thought process of a Mathematician.



Random Thinking And Its Limitations

- Many people may approach the "Game of 24" with random operations. For Example:

Random Approach

- Chose 14, 3 and multiply them

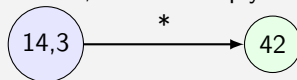


Random Thinking And Its Limitations

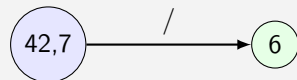
- Many people may approach the "Game of 24" with random operations. For Example:

Random Approach

- Chose 14, 3 and multiply them



- Now divide this 42 by 7

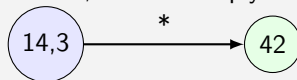


Random Thinking And Its Limitations

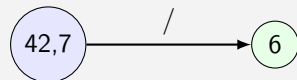
- Many people may approach the "Game of 24" with random operations. For Example:

Random Approach

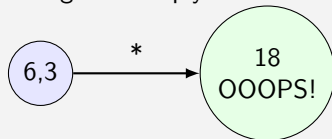
- Chose 14, 3 and multiply them



- Now divide this 42 by 7



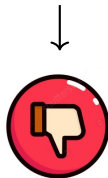
- Now again multiply this 6 with 3



Random Thinking and Its Limitations

From the previous slide, we see that a random approach to problem-solving is not very effective. Therefore, we propose:

- While random guessing might occasionally work for humans, it is inefficient and unreliable for machines.



Random Thinking and Its Limitations

From the previous slide, we see that a random approach to problem-solving is not very effective. Therefore, we propose:

- While random guessing might occasionally work for humans, it is inefficient and unreliable for machines.
- Instead, we will adopt the structured and systematic thought process of a mathematician.

How Mathematicians Solve Problems

How a mathematician will solve "Game of 24"! What are the differences?



How Mathematicians Solve Problems

How a mathematician will solve "Game of 24"! What are the differences?

A mathematician solves problems systematically:

- First He will try to sort the inputs.

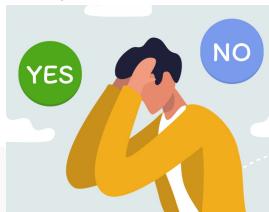


How Mathematicians Solve Problems

How a mathematician will solve "Game of 24"! What are the differences?

A mathematician solves problems systematically:

- First He will try to sort the inputs.
- Evaluate whether it is even possible to reach 24.



A mathematician solves problems systematically:

-

How Mathematicians Solve Problems

How a mathematician will solve "Game of 24"! What are the differences?

A mathematician solves problems systematically:

- First He will try to sort the inputs.
- Evaluate whether it is even possible to reach 24.
- Explore logical paths using operations and combinations.
- Refine results to find the solution.

Step 1: Sorting and Initial Evaluation

- Sort the inputs: 3, 5, 7, 14.

Step 1: Sorting and Initial Evaluation

- Sort the inputs: 3, 5, 7, 14.
- Check if any number is greater than 24 (None here).

Step 1: Sorting and Initial Evaluation

- Sort the inputs: 3, 5, 7, 14.
- Check if any number is greater than 24 (None here).
- Multiply all numbers: $3 \times 5 \times 7 \times 14 = 1470$.

Step 1: Sorting and Initial Evaluation

- Sort the inputs: 3, 5, 7, 14.
- Check if any number is greater than 24 (None here).
- Multiply all numbers: $3 \times 5 \times 7 \times 14 = 1470$.
- Since $1470 > 24$, a solution is possible.

Step 2: Finding the Largest Number And Possible Solution

- The largest number is 14.

Step 2: Finding the Largest Number And Possible Solution

- The largest number is 14.
- Check if the remaining numbers (3, 5, 7) can form 10: Not possible.

Step 2: Finding the Largest Number And Possible Solution

- The largest number is 14.
- Check if the remaining numbers (3, 5, 7) can form 10: Not possible.
- Move to the next step by exploring pairs.

Step 3: Pairing and Operations

Explore all pairs:

$$(3, 5) \rightarrow 3 + 5 = 8, 3 \times 5 = 15, 3 - 5 = -2, 3/5 = 0.6$$

$$(5, 7) \rightarrow 5 + 7 = 12, 5 \times 7 = 35, \dots$$

... (Other pairs similarly explored)

Step 3: Pairing and Operations

Explore all pairs:

$$(3, 5) \rightarrow 3 + 5 = 8, 3 \times 5 = 15, 3 - 5 = -2, 3/5 = 0.6$$

$$(5, 7) \rightarrow 5 + 7 = 12, 5 \times 7 = 35, \dots$$

... (Other pairs similarly explored)

Highlight: $5 \times 7 = 35$, which is a promising result.

Step 3: Pairing and Operations

Explore all pairs:

$$(3, 5) \rightarrow 3 + 5 = 8, 3 \times 5 = 15, 3 - 5 = -2, 3/5 = 0.6$$

$$(5, 7) \rightarrow 5 + 7 = 12, 5 \times 7 = 35, \dots$$

\dots (Other pairs similarly explored)

Highlight: $5 \times 7 = 35$, which is a promising result.

The remaining digits are 3 and 14 .

Step 4: Finding the Missing Subtractor

- From 35, He needs to subtract 11 to reach 24.

Step 4: Finding the Missing Subtractor

- From 35, He needs to subtract 11 to reach 24.
- He checks the remaining numbers: $14 - 3 = 11$, the desired subtractor.

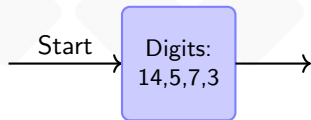
$$35 - 11 = 24$$

PROBLEM SOLVED !!!!!!!!!!!!!!!

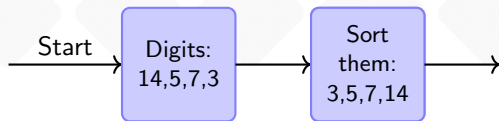
Bringing All Steps Together

Start →

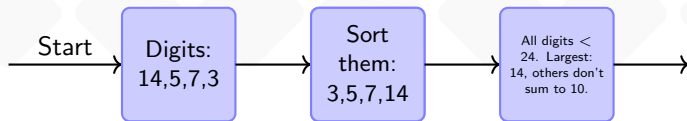
Bringing All Steps Together



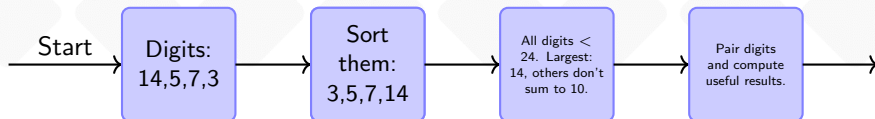
Bringing All Steps Together



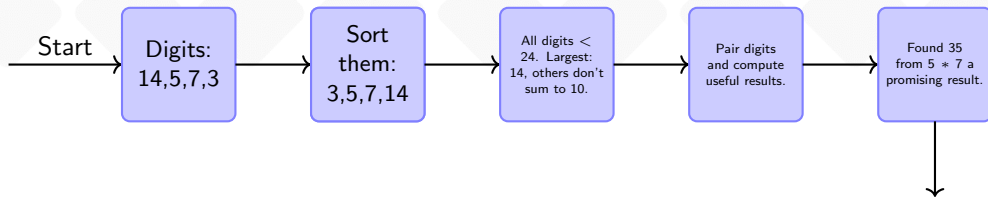
Bringing All Steps Together



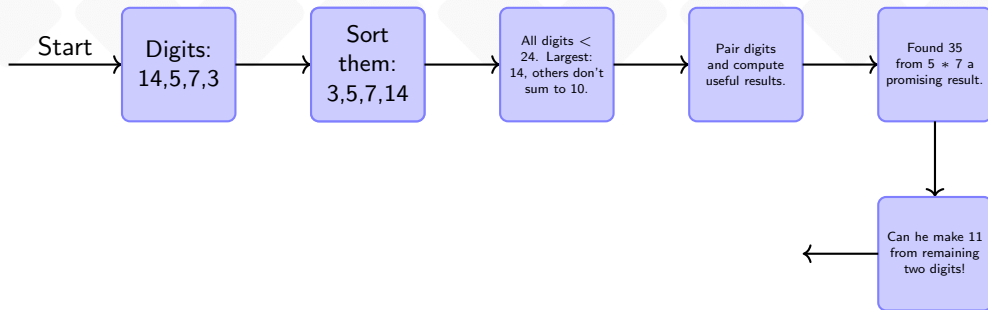
Bringing All Steps Together



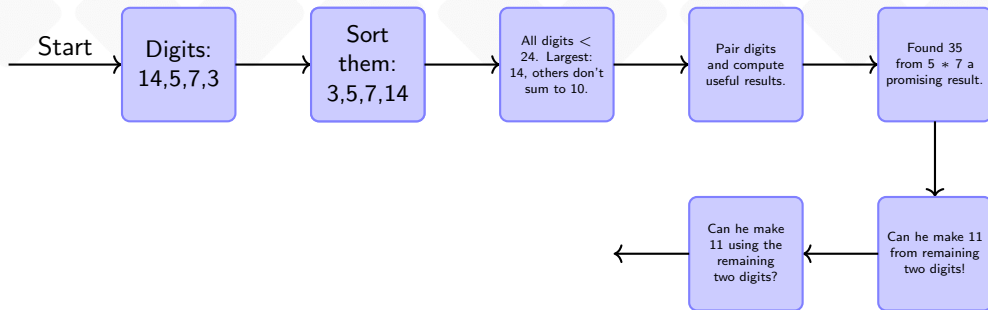
Bringing All Steps Together



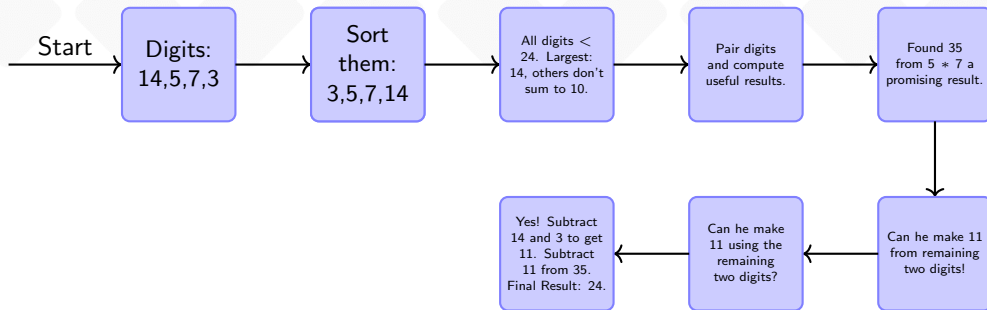
Bringing All Steps Together



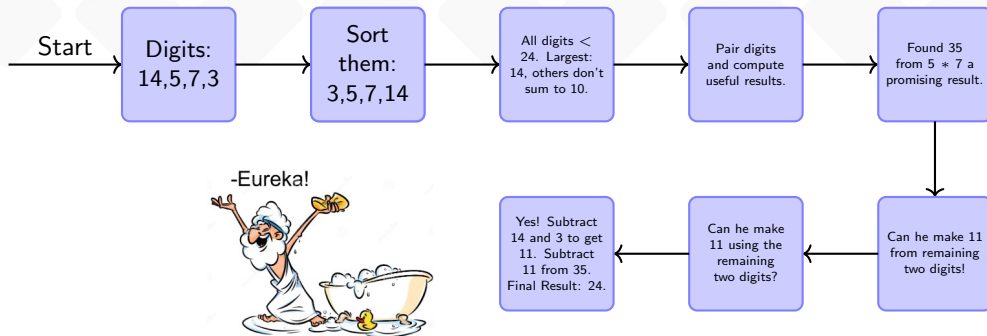
Bringing All Steps Together



Bringing All Steps Together



Bringing All Steps Together



Solution: $5 * 7 - 14 + 3 = 24$

Advantages of Structured Thinking in Problem Solving

- Structured reasoning eliminates random guesses and focuses on promising paths.

Advantages of Structured Thinking in Problem Solving

- Structured reasoning eliminates random guesses and focuses on promising paths.
- Logical exploration ensures efficiency and clarity in problem-solving.

Advantages of Structured Thinking in Problem Solving

- Structured reasoning eliminates random guesses and focuses on promising paths.
- Logical exploration ensures efficiency and clarity in problem-solving.
- This approach is not just for puzzles but applies to many areas of reasoning.

Modeling Human Problem-Solving in Language Models

Steps To solve the problem using AI tools using ToT prompting

- Problem Decomposition.

Modeling Human Problem-Solving in Language Models

Steps To solve the problem using AI tools using ToT prompting

- Problem Decomposition.
- Thought Generation.

Modeling Human Problem-Solving in Language Models

Steps To solve the problem using AI tools using ToT prompting

- Problem Decomposition.
- Thought Generation.
 - State Evaluation.

Modeling Human Problem-Solving in Language Models

Steps To solve the problem using AI tools using ToT prompting

- Problem Decomposition.
- Thought Generation.
 - State Evaluation.
 - Decision Making.

Step 1: Problem Decomposition

Language Models systematically break the problem into smaller parts:

- **Input Analysis:** Identify the components:

Numbers: 3, 5, 7, 14

Target: 24

Step 1: Problem Decomposition

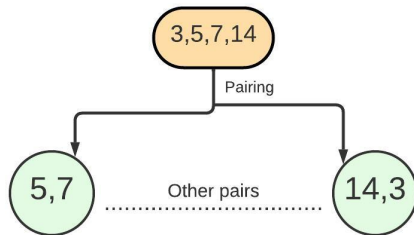
Language Models systematically break the problem into smaller parts:

- **Input Analysis:** Identify the components:

Numbers: 3, 5, 7, 14

Target: 24

- In this step pairs of inputs are made.
- **Pair the Inputs:**



Step 2: Thought Generation

Language Models explore multiple paths to generate solutions:

- **Objective:** Systematically compute results from the pairs generated in the decomposition step and apply operations to approach the target 24.

Step 2: Thought Generation

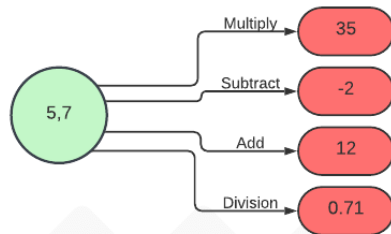
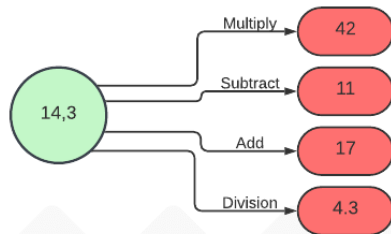
Language Models explore multiple paths to generate solutions:

- **Objective:** Systematically compute results from the pairs generated in the decomposition step and apply operations to approach the target 24.
- From the decomposition step, pairs such as (5, 7) and (14, 3) are identified.

Step 2: Thought Generation

Language Models explore multiple paths to generate solutions:

- **Objective:** Systematically compute results from the pairs generated in the decomposition step and apply operations to approach the target 24.
- From the decomposition step, pairs such as (5, 7) and (14, 3) are identified.
- Different arithmetic operations are performed on these pairs to generate intermediate values.
- For example:



Intermediate Values

Game of 24 Prompt

Conceptual Stages or Processes:

- **Propose Prompt**
- **Value Prompt**

Game of 24 Prompt

Propose Prompt

User: Game of 24 is a mathematical reasoning challenge, where the goal is to use 4 numbers and basic arithmetic operations (+,-,*,/) to obtain 24. input: 3,5,7,14 give me all the possible next step that will left me with 3 input value

Assistant: To determine all possible next steps in the "Game of 24" for the input numbers 3,5,7,14, we combine any two numbers using the basic arithmetic operations (+, -, *, /), provided division results in an integer. This reduces the number of inputs from 4 to 3.

Steps:

- Pair any two numbers (e.g., 3 and 5).
- Apply each valid operation to create a new number.
- Replace the pair with the result, leaving three numbers.

Game of 24 Prompt

Propose Prompt Continue

Combinations and Results

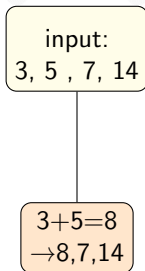
1. Pair 3 and 5:

- $3 + 5 = 8 \rightarrow$ Remaining: 8, 7, 14
- $3 - 5 = -2 \rightarrow$ Remaining: -2, 7, 14
- $5 - 3 = 2 \rightarrow$ Remaining: 2, 7, 14
- $3 \times 5 = 15 \rightarrow$ Remaining: 15, 7, 14

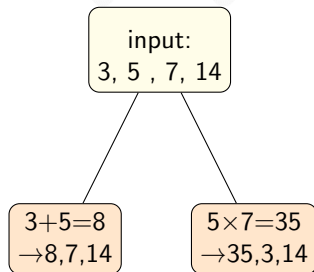
Game of 24

input:
3, 5 , 7, 14

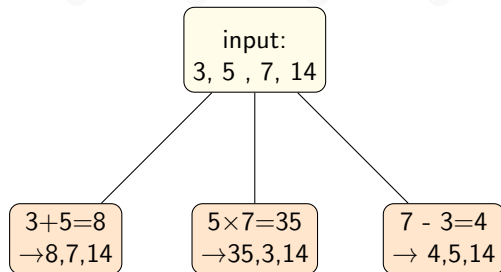
Game of 24



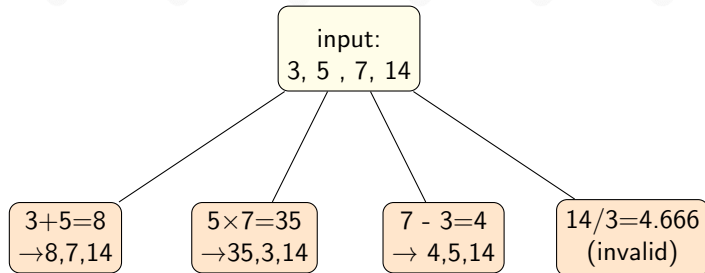
Game of 24



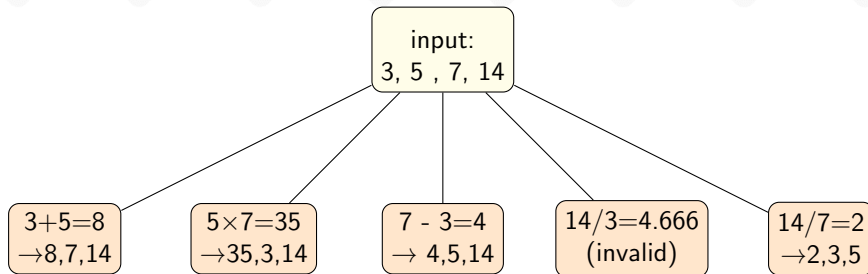
Game of 24



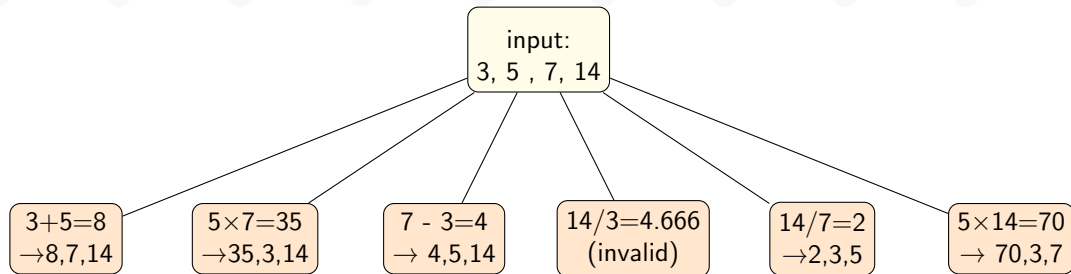
Game of 24



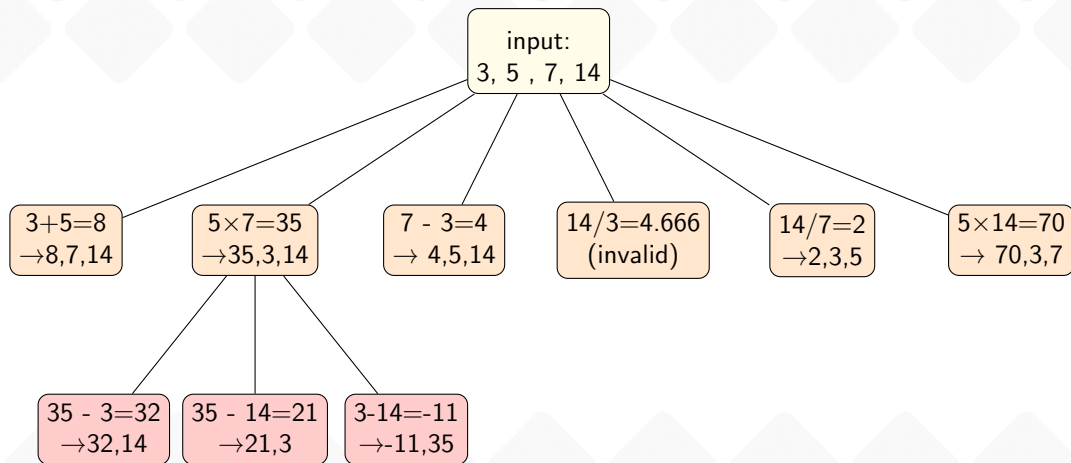
Game of 24



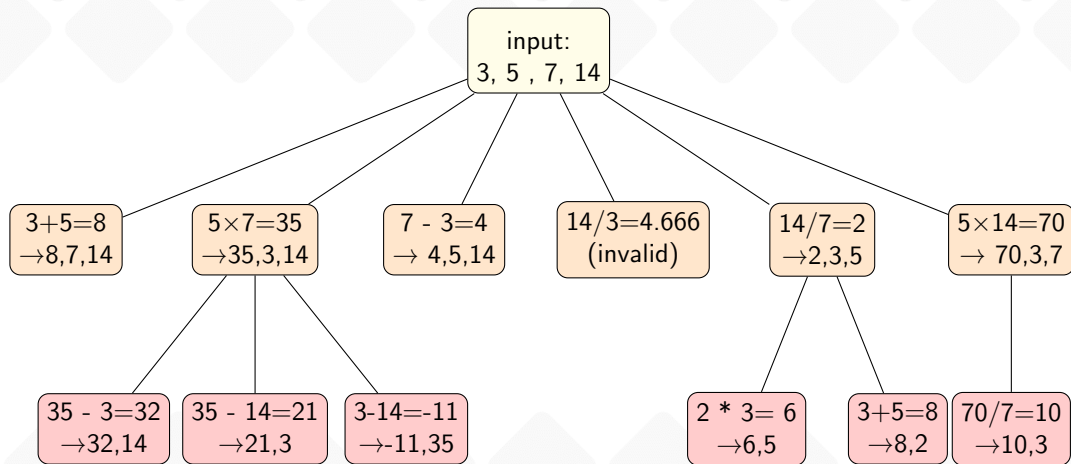
Game of 24



Game of 24



Game of 24



Game of 24 Prompt

Value Prompt

User: Evaluate if given numbers can reach 24 (sure/likely/impossible)
input: 21,3

Assistant: To evaluate whether the numbers 21 and 3 can be combined using basic arithmetic operations (+, -, *, /) to reach 24, we need to explore all possible combinations.

1. Addition:

- $21 + 3 = 24$
- **Success!**

Game of 24 Prompt

Value Prompt Continue

2. Subtraction:

- $21 - 3 = 18$ (not 24)
- $3 - 21 = -18$ (not 24)

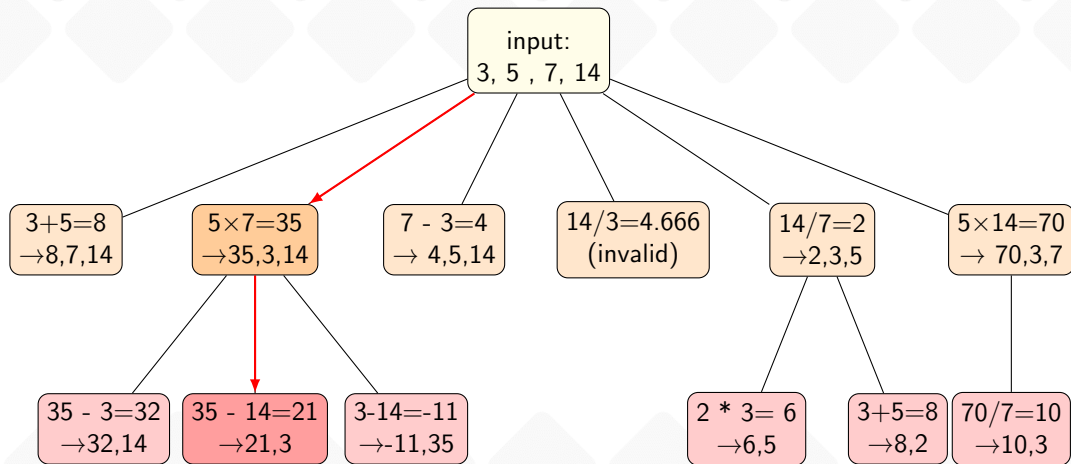
3. Multiplication:

- $21 \times 3 = 63$ (not 24)

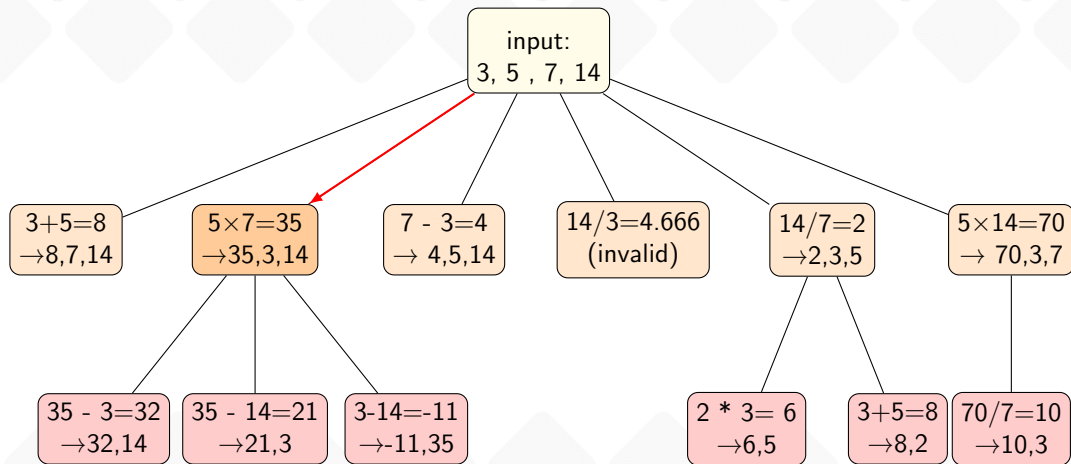
4. Division:

- $21 \div 3 = 7$ (not 24)
- $3 \div 21 = 0.142857$ (not 24)

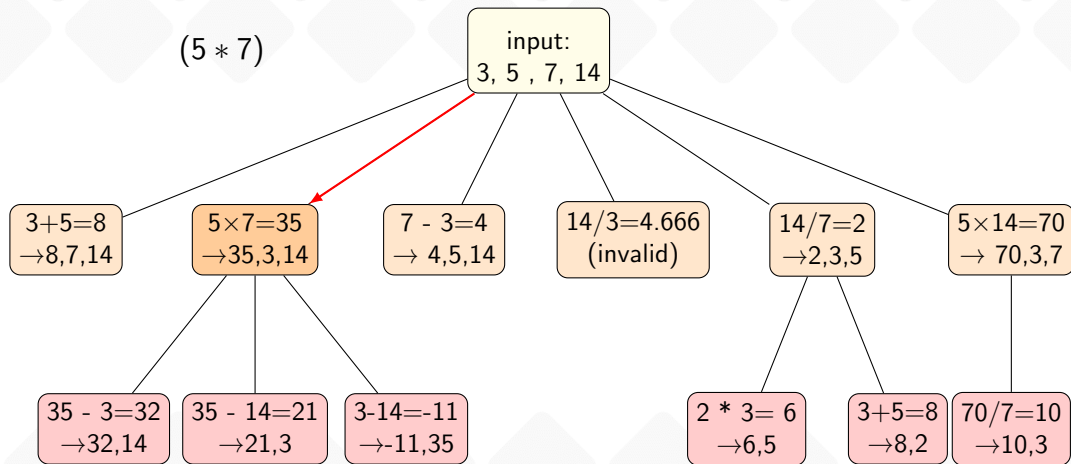
Game of 24



Game of 24

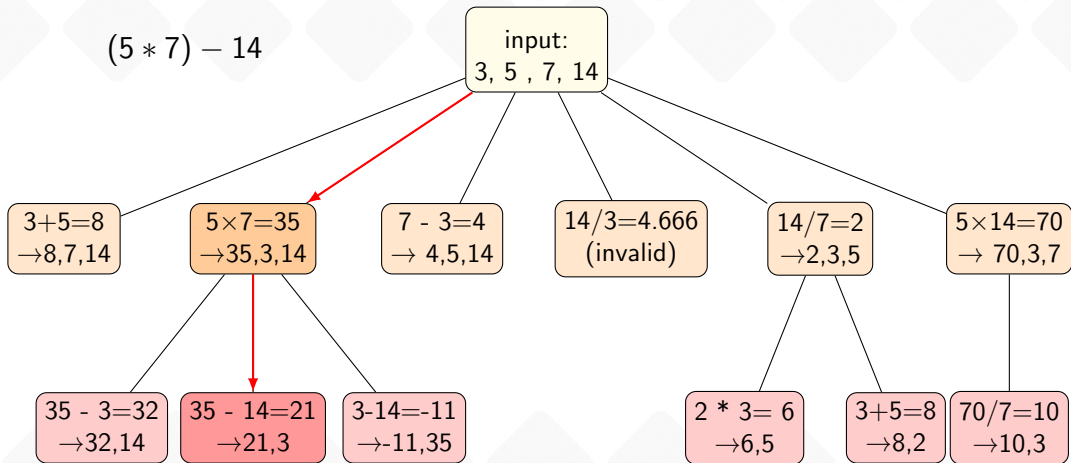


Game of 24



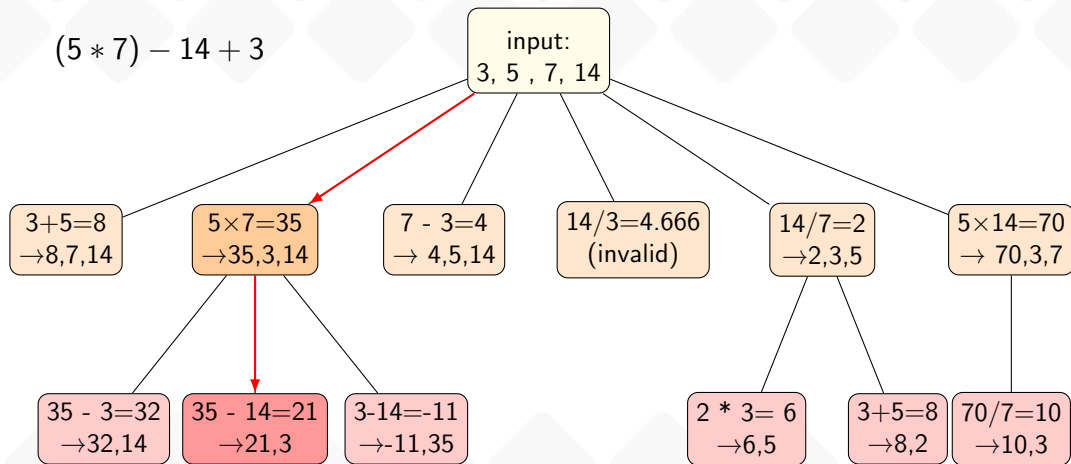
Game of 24

$$(5 * 7) - 14$$



Game of 24

$$(5 * 7) - 14 + 3$$



ToT in a game of 24

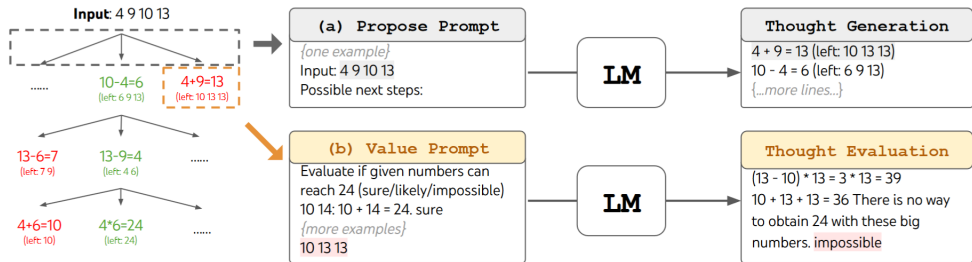


Figure 1 – ToT in a game of 24. The LM is prompted for (a) thought generation and (b) valuation..

limitations

Resource Intensive:

Game of 24	Generate/Prompt tokens	Cost per case	Success
IO (best of 100)	1.8k / 1.0k	\$0.13	33%
CoT (best of 100)	6.7k / 2.2k	\$0.47	49%
ToT	5.5k / 1.4k	\$0.74	74%

Tabela 1 – Cost analysis on Game of 24..

Limited Applicability:Limited to tasks requiring structured and logical problem-solving.

Capabilities of ToT Framework

- **Enhanced Decision-Making.**
- **Applied primarily to reasoning and search problems.**

Thank You!