

Prompting to Reason Deductively

Dorota Nowak

Jagiellonian University

dorota.iza.nowak@student.uj.edu.pl

Mateusz Pach

Jagiellonian University

mateusz.pach@student.uj.edu.pl

Michał Wronka

Jagiellonian University

michal.wronka@student.uj.edu.pl

Andrii Krutsylo*

Jagiellonian University

andrii.krutsylo@uj.edu.pl

Paweł Budzianowski*

Jagiellonian University

pawel.budzianowski@ii.uj.edu.pl

Abstract

This paper investigates the effectiveness of different prompting strategies in improving the mathematical reasoning abilities of large language models (LLMs) based on transformer architecture. We propose and evaluate various prompting strategies, including chain-of-thought approaches. Experimental results demonstrate that more complex prompts do not necessarily lead to better performance and that the Complex-CoT strategy achieves the highest accuracy. The findings highlight the importance of prompt design in enhancing mathematical reasoning capabilities in LLMs.

1 Introduction

The transformer architecture (Vaswani et al., 2017) allowed us to build *Large Language Models* (LLMs) that are extremely good at remembering a wide amount of knowledge (Brown et al., 2020; OpenAI, 2023; Touvron et al., 2023; Lewis et al., 2020; Chowdhery et al., 2022). In particular, it allows them to perform limited logical and mathematical reasoning (OpenAI, 2023). However, they struggle to extrapolate it to larger numbers (Pal and Baral, 2021).

To counteract this, several approaches (Zhang, 2023; Jie et al., 2022) tried to reduce the transformer’s role to producing a general expression by adapting the architecture specifically for this task. Such expressions can then be deterministically evaluated to obtain the needed values. However, recent experiments show that these approaches applied on smaller LLMs like RoBERTa (Liu et al., 2019) perform worse than even the largest LLMs such as PALM (Chowdhery et al., 2022) used directly (Kojima et al., 2023). The difference is even greater

when we improve query prompts with techniques like few-shot, chain-of-thought (Wei et al., 2022; Kojima et al., 2023), self-consistency (Wang et al., 2022), or progressive-hint prompting (Zheng et al., 2023).

Inspired by these ideas, we raise the following question: Can we improve on existing prompting strategies by combining the observations found in some of them with the design idea of previous state-of-the-art network architecture, which finds general expressions (Zhang, 2023; Jie et al., 2022)?

2 Related work

2.1 Math word problem

Math word problem (MWP) is the task of answering a mathematical question described in the natural language (Bobrow, 1964). The authors argue that the use of formal programming languages and mathematical notation can be a barrier to entry for users who are not familiar with these notations. They propose using natural language as input to allow users to interact with the system in a more intuitive way. The developed system (STUDENT) was able to parse the natural language input, convert it into a formal representation, and use symbolic manipulation techniques to solve the problem. The authors reported promising results and since then many new approaches have been created.

2.2 LLMs

It has been shown that large language models can achieve exceptional results in different natural language tasks. However, they can struggle with tasks where mathematical reasoning is required. The authors of (Pal and Baral, 2021) test this numeracy learning ability of a text-to-text transfer learning

*Project Supervisors

generative model. The model works quite well in a setting where the training and testing are performed on the same numerical range of data but struggles when training is done on lower and testing on a higher range of data. The model proposed in (Chowdhery et al., 2022) achieved state-of-the-art few-shot results across hundreds of natural language and mathematical reasoning tasks. The proposed transformed-based model is pre-trained on a collection of 780 billion tokens that cover diverse natural language tasks and applications.

2.3 Task-specific approaches

Zhang (2023) and Jie et al. (2022) aim to obtain the answer by producing a complete computation tree.

They start by extracting a list of quantities present in the given MWP instance and converting them to a representation that is later appended to the model input.

Then, they use an NLP transformer (*Reasoner*) to produce combinations of two values (chosen by the model from the representation) with primitive binary operators (also chosen by the model). Afterward, the *Terminator* network decides if the instance’s answer was produced, based on a score attached to each combination (obtained from a separate *Scorer* network).

If the answer is decided to not be present yet, the *Rationalizer* network extends the representation with the highest score combinations produced in the current step. Then the procedure is repeated.

If the model performs more than T_{max} iterations without producing the output, the highest-scoring combination is returned instead. The schema of the (Zhang, 2023) model can be seen in Figure 1.

2.4 Prompting strategies

Several simple, but effective strategies of prompting have been proposed so far. Wei et al. (2022) add intermediate steps of reasoning to the solutions in a few-shot setting. Kojima et al. (2023) basically append the sentence “Let’s think step by step.” Wang et al. (2022) sample a few LLM answers and then choose the most consistent one. Zheng et al. (2023) repeat the same question to the LLM, each time attaching the previous results as hints.

3 Experiments

We start by defining two existing strategies (Simple and Complex-CoT), followed by three novel ones

(Complex-Deductive, Complex-Deductive-Natural-Language, and Complex-Simplify). We then perform experiments in two setups described later^{*}. Additionally, we propose and test an interactive way of prompting that simulates the method of Zhang (2023) and Jie et al. (2022).

3.1 Strategies

3.1.1 Simple

Our baseline strategy outputs a single sentence: “The answer is x ” (see A.1.1).

3.1.2 Complex-CoT

Similarly, as Wei et al. (2022) we pick 3 of the most complex tasks in the training dataset and require the Chain-of-Thought answer (see A.1.2). It is supposed to help the model deduce in a step-by-step manner. Like in all of the following strategies, the name prefix Complex comes from picking the hardest tasks instead of random ones.

3.1.3 Complex-Deductive

This strategy uses mathematical variables and equations as a systematic approach to reach the ultimate solution. First, we create a list of all available variables and the unknown and after that, we sequentially expand the variables list until we find the answer (see A.1.3).

3.1.4 Complex-Deductive-Natural-Language

In this version, we adopt a more intuitive approach, and instead of using mathematical variables and equations, we conduct all the calculations using natural language. This approach was supposed to make the complex computations more understandable (see A.1.4).

3.1.5 Complex-Simplify

In this method, we started by simplifying the original question while keeping all the important information. We thought that by rephrasing the question, it might become easier for the model to understand (see A.1.5).

3.2 Few-shot

In our initial experiment, we tried various prompting strategies (described in the previous section) on the GPT-3.5-turbo model. We used the GSM8K (Cobbe et al., 2021) dataset to conduct this study and run the first 100 questions from the testing dataset. Then, we analyzed the model’s responses

^{*}<https://github.com/mateuszpach/transformers-for-mwp>

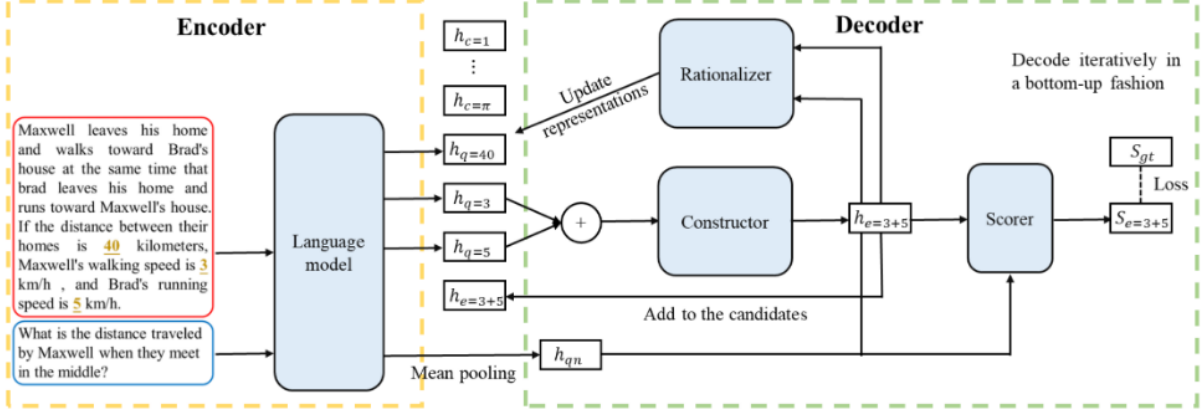


Figure 1: Framework overview of the model proposed by Zhang (2023).

Simple	Complex-CoT	Complex-Deductive	Complex-Deductive-Natural-Language	Complex-Simplify
0.32	0.76	0.63	0.67	0.73

Table 1: GSM8K accuracy on GPT-3.5-turbo with few-shot prompting.

	Baseline CoT	Complex-CoT	Complex-Deductive	Complex-Deductive-Natural-Language	Complex-Simplify
flan-t5-base	0.024	0.025	0.019	0.024	0.017
flan-t5-large	0.048	0.044	0.024	0.033	0.046

Table 2: GSM8K accuracy on fine-tuned Flan-T5.

to determine the number of accurate answers returned. We provide the results in Table 1.

3.3 Fine-tuning

We tried to fine-tune Google’s *Flan-T5* (Chung et al., 2022) on our strategies. As training data for each strategy, we used a collection of answers generated by the previous experiment (only those producing correct outputs). We fine-tuned Flan-T5-Base and Flan-T5-Large (due to hardware limitations we were unable to attempt fine-tuning larger models).

The baseline was fine-tuned on original answers from GSM8K, which are organic, hand-crafted CoT answers with equations. Accuracy is measured as before by comparing the last numeric value present in the answer with the ground truth answer. We present our results in the Table 2.

3.4 Interactive zero-shot

Lastly, we tried to simulate iterative reasoning from (Zhang, 2023) on GPT-3.5-turbo, with a strategy that initially asked the model for a list of variables present in the question and then requested it to extend the list with new variables built from those already present in the list. After each iteration, the model was asked if the answer is present in the

list and asked to return it if so. This schema can be formulated as finite state automata responding to GPT outputs. All such strategies plateaued at around 31% accuracy.

4 Discussion

In the few-shot experiments, we can see that the simplest prompt produced the worst results and that improving the complexity of the prompt improves the output. Yet again, the most complicated prompt does not always yield the best results. The second worse is the Complex-Deductive prompt, which uses mathematical expressions and we can assume that the model gains from providing examples using natural language. The best accuracy was achieved by the Complex-CoT prompt.

The fine tuning experiment shows that, Flan-T5 does not perform well on our task with best achieved accuracy around only 4.8%. Furthermore, the most important parameter in this experiment is size of the network, with all strategies on larger model achieving 1.5–2.8 times better scores. What is more, even though none of our strategies significantly outperformed the Baseline CoT, we proved that training data generated by GPT can be as good as hand-crafted. Additionally, we conclude that Complex-CoT, preforms comparably with baseline

regardless of model size, while Complex-Simplify achieves significant improvement on larger NL models; we speculate that NL models need a set threshold of size below which CoT is the best strategy and only once a model is expressive enough to perform this task it can further benefit from additionally requesting it to rephrase/simplify the question. Finally, deductive strategies by introducing additional requirements like using variable substitution or rigid "if x then y " sentence structure seem to obfuscate the task at hand; this is likely caused because NL models at the end of the day are not reasoning in any human-like capacity, but simply predict the next word given the previous ones, so allowing the model more freedom in sentence structure will likely lead to better result, especially when it comes to math equations, where model might remember that " $5 * 6 =$ " is followed by "30", but is less likely to finish " $(a + b)/c =$ " with correct substitution of values.

Experiment with interactive prompting showed that such strategy, did not outperform directly asking GPT for the answer. Despite being able to mimic the steps of the algorithms such as proposed by Zhang (2023) or Jie et al. (2022), the model fails either to follow them or on calculations.

5 Conclusions

It is challenging to definitively determine the optimal prompting strategy for math problem solving and further research is needed in this area. It is important to note that different models may yield diverse results, making it necessary to explore multiple models to gain a comprehensive understanding of the most effective promoting strategies.

References

- Daniel Bobrow. 1964. Natural language input for a computer problem solving system.
- Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. [Language models are few-shot learners](#).
- Aakanksha Chowdhery, Sharan Narang, Jacob Devlin, Maarten Bosma, Gaurav Mishra, Adam Roberts, Paul Barham, Hyung Won Chung, Charles Sutton, Sebastian Gehrmann, Parker Schuh, Kensen Shi, Sasha Tsvyashchenko, Joshua Maynez, Abhishek Rao, Parker Barnes, Yi Tay, Noam Shazeer, Vinodkumar Prabhakaran, Emily Reif, Nan Du, Ben Hutchinson, Reiner Pope, James Bradbury, Jacob Austin, Michael Isard, Guy Gur-Ari, Pengcheng Yin, Toju Duke, Anselm Levskaya, Sanjay Ghemawat, Sunipa Dev, Henryk Michalewski, Xavier Garcia, Vedant Misra, Kevin Robinson, Liam Fedus, Denny Zhou, Daphne Ippolito, David Luan, Hyeontaek Lim, Barret Zoph, Alexander Spiridonov, Ryan Sepassi, David Dohan, Shivani Agrawal, Mark Omernick, Andrew M. Dai, Thanumalayan Sankaranarayanan Pillai, Marie Pellat, Aitor Lewkowycz, Erica Moreira, Rewon Child, Oleksandr Polozov, Katherine Lee, Zongwei Zhou, Xuezhi Wang, Brennan Saeta, Mark Diaz, Orhan Firat, Michele Catasta, Jason Wei, Kathy Meier-Hellstern, Douglas Eck, Jeff Dean, Slav Petrov, and Noah Fiedel. 2022. [Palm: Scaling language modeling with pathways](#).
- Hyung Won Chung, Le Hou, Shayne Longpre, Barret Zoph, Yi Tay, William Fedus, Yunxuan Li, Xuezhi Wang, Mostafa Dehghani, Siddhartha Brahma, Albert Webson, Shixiang Shane Gu, Zhuyun Dai, Mirac Suzgun, Xinyun Chen, Aakanksha Chowdhery, Alex Castro-Ros, Marie Pellat, Kevin Robinson, Dasha Valter, Sharan Narang, Gaurav Mishra, Adams Yu, Vincent Zhao, Yanping Huang, Andrew Dai, Hongkun Yu, Slav Petrov, Ed H. Chi, Jeff Dean, Jacob Devlin, Adam Roberts, Denny Zhou, Quoc V. Le, and Jason Wei. 2022. [Scaling instruction-finetuned language models](#).
- Karl Cobbe, Vineet Kosaraju, Mohammad Bavarian, Mark Chen, Heewoo Jun, Lukasz Kaiser, Matthias Plappert, Jerry Tworek, Jacob Hilton, Reiichiro Nakano, Christopher Hesse, and John Schulman. 2021. Training verifiers to solve math word problems. *arXiv preprint arXiv:2110.14168*.
- Zhanming Jie, Jierui Li, and Wei Lu. 2022. Learning to reason deductively: Math word problem solving as complex relation extraction. In *Annual Meeting of the Association for Computational Linguistics*.
- Takeshi Kojima, Shixiang Shane Gu, Machel Reid, Yutaka Matsuo, and Yusuke Iwasawa. 2023. [Large language models are zero-shot reasoners](#).
- Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer. 2020. [BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension](#). In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pages 7871–7880, Online. Association for Computational Linguistics.
- Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019.

Roberta: A robustly optimized bert pretraining approach.

OpenAI. 2023. Gpt-4 technical report.

Kuntal Kumar Pal and Chitta Baral. 2021. [Investigating numeracy learning ability of a text-to-text transfer model](#). In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3095–3101, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, Aurelien Rodriguez, Armand Joulin, Edouard Grave, and Guillaume Lample. 2023. [Llama: Open and efficient foundation language models](#).

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. [Attention is all you need](#).

Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Huai hsin Chi, and Denny Zhou. 2022. Self-consistency improves chain of thought reasoning in language models. *ArXiv*, abs/2203.11171.

Jason Wei, Xuezhi Wang, Dale Schuurmans, Maarten Bosma, Ed Huai hsin Chi, F. Xia, Quoc Le, and Denny Zhou. 2022. Chain of thought prompting elicits reasoning in large language models. *ArXiv*, abs/2201.11903.

Youyuan Zhang. 2023. Techniques to improve neural math word problem solvers. *ArXiv*, abs/2302.03145.

Chuanyang Zheng, Zhengying Liu, Enze Xie, Zhenguo Li, and Yu Li. 2023. [Progressive-hint prompting improves reasoning in large language models](#).

A Appendix

A.1 Few-shot prompts used in each strategy

A.1.1 Simple

Follow the examples and respond to the question in a similar manner. You can do all the reasoning you want, but your last sentence has to begin with 'The answer is '.

Q: There are 15 trees in the grove. Grove workers will plant trees in the grove today. After they are done, there will be 21 trees. How many trees did the grove workers plant today?

A: The answer is 6

Q: If there are 3 cars in the parking lot and 2 more cars arrive, how many cars are in the parking lot?

A: The answer is 5

Q: Olivia has \$23. She bought five bagels for \$3 each. How much money does she have left?

A: The answer is 8

A.1.2 Complex-CoT

Follow the examples and respond to the question in a similar manner. You can do all the reasoning you want, but your last sentence has to begin with 'The answer is '.

Q: Mark's basketball team scores 25 2 pointers, 8 3 pointers and 10 free throws. Their opponents score double the 2 pointers but half the 3 pointers and free throws. What's the total number of points scored by both teams added together?

A: Let's think step by step Mark's team scores 25 2 pointers, meaning they scored $25 \times 2 = 50$ points in 2 pointers. His team also scores 6 3 pointers, meaning they scored $8 \times 3 = 24$ points in 3 pointers They scored 10 free throws, and free throws count as one point so they scored $10 \times 1 = 10$ points in free throws. All together his team scored $50 + 24 + 10 = 84$ points Mark's opponents scored double his team's number of 2 pointers, meaning they scored $50 \times 2 = 100$ points in 2 pointers. His opponents scored half his team's number of 3 pointers, meaning they scored $24 / 2 =$

12 points in 3 pointers. They also scored half Mark's team's points in free throws, meaning they scored $10/2=5$ points in free throws. All together Mark's opponents scored $100+12+5=117$ points The total score for the game is both team's scores added together, so it is $84+117=201$ points The answer is 201

Q: Bella has two times as many marbles as frisbees. She also has 20 more frisbees than deck cards. If she buys $2/5$ times more of each item, what would be the total number of the items she will have if she currently has 60 marbles? A: Let's think step by step When Bella buys $2/5$ times more marbles, she'll have increased the number of marbles by $2/5 \times 60 = 24$ The total number of marbles she'll have is $60+24 = 84$ If Bella currently has 60 marbles, and she has two times as many marbles as frisbees, she has $60/2 = 30$ frisbees. If Bella buys $2/5$ times more frisbees, she'll have $2/5 \times 30 = 12$ more frisbees. The total number of frisbees she'll have will increase to $30+12 = 42$ Bella also has 20 more frisbees than deck cards, meaning she has $30-20 = 10$ deck cards If she buys $2/5$ times more deck cards, she'll have $2/5 \times 10 = 4$ more deck cards. The total number of deck cards she'll have is $10+4 = 14$ Together, Bella will have a total of $14+42+84 = 140$ items The answer is 140

Question: A group of 4 fruit baskets contains 9 apples, 15 oranges, and 14 bananas in the first three baskets and 2 less of each fruit in the fourth basket. How many fruits are there?

A: Let's think step by step For the first three baskets, the number of apples and oranges in one basket is $9+15=24$ In total, together with bananas, the number of fruits in one basket is $24+14=38$ for the first three baskets. Since there are three baskets each having 38 fruits, there are $3 \times 38=114$ fruits in the first three baskets. The number of apples in the fourth basket is $9-2=7$ There are also $15-2=13$ oranges in the fourth basket The combined number of oranges and apples in the fourth basket is $13+7=20$ The fourth basket also contains $14-2=12$ bananas. In

total, the fourth basket has $20+12=32$ fruits. The four baskets together have $32+114=146$ fruits. The answer is 146

A.1.3 Complex-Deductive

Follow the examples and respond to the question in a similar manner. You can do all the reasoning you want, but your last sentence has to begin with 'The answer is '.

Q: Mark's basketball team scores 25 2 pointers, 8 3 pointers and 10 free throws. Their opponents score double the 2 pointers but half the 3 pointers and free throws. What's the total number of points scored by both teams added together? A: The variables in the task are: $x = 25$ - the number of 2 pointers scored by Mark's team $y = 8$ - the number of 3 pointers scored by Mark's team $z = 10$ - the number of free throws scored by Mark's team $m = 2$ - what part of the 2 pointers scored by Mark's team did his opponents score $n = 0.5$ - what part of the 3 pointers scored by Mark's team did his opponents score $o = 0.5$ - what part of the free throws scored by Mark's team did his opponents score The goal is to find: $p = ?$ - the total number of points scored by both teams added together Let's sequentially expand the variables list until we find the answer. $x * 2 = 25 * 2 = 50$ - the number of points scored by Mark's team in 2 pointers $y * 3 = 8 * 3 = 24$ - the number of points scored by Mark's team in 3 pointers $z * 1 = 10 * 1 = 10$ - the number of points scored by Mark's team in free throws $(x * 2) * m = 50 * 2 = 100$ - the number of points scored by Mark's opponents in 2 pointers $(y * 3) * n = 24 * 0.5 = 12$ - the number of points scored by Mark's opponents in 3 pointers $(z * 1) * o = 10 * 0.5 = 5$ - the number of points scored by Mark's opponents in free throws $(x * 2) + (y * 3) + (z * 1) = 50 + 24 + 10 = 84$ - the total number of points scored by Mark's team $((x * 2) * m) + ((y * 3) * n) + ((z * 1) * o) = 100 + 12 + 5 = 117$ - the total number of points scored by Mark's opponents $((x * 2) + (y * 3) + (z * 1)) + (((x * 2) * m) + ((y * 3) * n) + ((z * 1) * o)) = 84 + 117 =$

201 - the total number of points scored by both teams added together The answer is 201

Q: Bella has two times as many marbles as frisbees. She also has 20 more frisbees than deck cards. If she buys $\frac{2}{5}$ times more of each item, what would be the total number of the items she will have if she currently has 60 marbles?

A: The variables in the task are: $x = 2$ - the number of marbles Bella has for each frisbee $y = 20$ - the number of frisbees Bella has more than deck cards $z = 0.4$ - the number of items Bella buys more per each owned item $k = 60$ - the number of marbles Bella currently has The goal is to find: $p = ?$ - the total number of items Bella will have if after buying more items Let's sequentially expand the variables list until we find the answer. $k / x = 60 / 2 = 30$ - the number of frisbees Bella has $k / x - y = 30 - 20 = 10$ - the number of deck cards Bella has $k + k / x + k / x - y = 60 + 30 + 10 = 100$ - the total number of items Bella currently has $(k * z + (k / x) * z + (k / x - y)) * z = (60 * 0.4 + (30 * 0.4) + (10 * 0.4)) * 0.4 = 24 + 12 + 4 = 40$ - the total number of items Bella will buy $(k + k / x + k / x - y) + ((k * z + (k / x) * z + (k / x - y)) * z) = 100 + 40 = 140$ - the total number of items Bella will have The answer is 140

Question: A group of 4 fruit baskets contains 9 apples, 15 oranges, and 14 bananas in the first three baskets and 2 less of each fruit in the fourth basket. How many fruits are there?

A: The variables in the task are: $g = 3$ - the number of baskets containing a apples, o oranges, and b bananas $a = 9$ - the number of apples in the first g baskets $o = 15$ - the number of oranges in the first g baskets $b = 14$ - the number of bananas in the first g baskets $c = 2$ - how many less of each fruit are in the last basket The goal is to find: $p = ?$ - the total number of fruits in the four baskets Let's sequentially expand the variables list until we find the answer. $a + o + b = 9 + 15 + 14 = 38$ - the number of fruits in each of the first g baskets $(a + o + b)$

$* g = 38 * 3 = 114$ - the number of fruits in the first g baskets $a - c = 9 - 2 = 7$ - the number of apples in the last basket $o - c = 15 - 2 = 13$ - the number of oranges in the last basket $b - c = 14 - 2 = 12$ - the number of bananas in the last basket $a - c + o - c + b - c = 7 + 13 + 12 = 32$ - the number of fruits in the last basket $(a + o + b) * g + (a - c + o - c + b - c) = 114 + 32 = 146$ - the total number of fruits in the four baskets The answer is 146

A.1.4 Complex-Deductive-Natural-Language

Follow the examples and respond to the question in a similar manner. You can do all the reasoning you want, but your last sentence has to begin with 'The answer is '.

Q: Mark's basketball team scores 25 2 pointers, 8 3 pointers and 10 free throws. Their opponents score double the 2 pointers but half the 3 pointers and free throws. What's the total number of points scored by both teams added together?

A: Let's think deductively If Mark's team scores 25 2 pointers, then he scores 50 points in 2 pointers. If Mark's team scores 8 3 pointers, then he scores 24 points in 3 pointers. If Mark's team scores 10 free throws, then he scores 10 points in free throws. If Mark's opponents score double the 2 pointers than Mark's team which scores 50 points in 2 pointers, then they score 100 points in 2 pointers. If Mark's opponents score half the 3 pointers than Mark's team which scores 24 points in 3 pointers, then they score 12 points in 3 pointers. If Mark's opponents score half the free throws than Mark's team which scores 10 points in free throws, then they score 5 points in free throws. If Mark's team scores 50 points in 2 pointers, 24 points in 3 pointers and 10 points in free throws, then they score 84 points in total. If Mark's opponents score 100 points in 2 pointers, 12 points in 3 pointers and 5 points in free throws, then they score 117 points in total. If Mark's team scores 84 points in total and Mark's opponents score 117 points in

total, then both teams score 201 points in total. The answer is 201

Q: Bella has two times as many marbles as frisbees. She also has 20 more frisbees than deck cards. If she buys $\frac{2}{5}$ times more of each item, what would be the total number of the items she will have if she currently has 60 marbles?

A: Let's think deductively If Bella has 60 marbles and two times as many marbles as frisbees, then she has 30 frisbees. If Bella has 30 frisbees and 20 more frisbees than deck cards, then she has 10 deck cards. If Bella has 60 marbles, 30 frisbees and 10 deck cards, then she has 100 items in total. If Bella buys $\frac{2}{5}$ times more of each item and now she has 100 items, then she buys 40 items in total. If Bella has 100 items in total and buys 40 items in total, then she has 140 items in total. The answer is 140

Question: A group of 4 fruit baskets contains 9 apples, 15 oranges, and 14 bananas in the first three baskets and 2 less of each fruit in the fourth basket. How many fruits are there?

A: Let's think deductively If there are 9 apples, 15 oranges, and 14 bananas in the first three baskets, then there are 38 fruits in each of the first three baskets. If there are 38 fruits in each of the first three baskets, then there are 114 fruits in the first three baskets. If there are 9 apples and 2 less of each fruit in the fourth basket, then there are 7 apples in the fourth basket. If there are 15 oranges and 2 less of each fruit in the fourth basket, then there are 13 oranges in the fourth basket. If there are 14 bananas and 2 less of each fruit in the fourth basket, then there are 12 bananas in the fourth basket. If there are 7 apples, 13 oranges, and 12 bananas in the fourth basket, then there are 32 fruits in the fourth basket. If there are 114 fruits in the first three baskets and 32 fruits in the fourth basket, then there are 146 fruits in total. The answer is 146

A.1.5 Complex-Simplify

Follow the examples and respond to the question in a similar manner. You can do all the reasoning you want, but your last sentence has to begin with 'The answer is '.

Q: Mark's basketball team scores 25 2 pointers, 8 3 pointers and 10 free throws. Their opponents score double the 2 pointers but half the 3 pointers and free throws. What's the total number of points scored by both teams added together?

A: Let's simplify the problem statement without skipping any important information Mark's basketball team scored 25 two-pointers, 8 three-pointers, and 10 free throws. Their opponents scored twice as many two-pointers but half as many three-pointers and free throws. What is the combined total of points scored by both teams? Let's think step by step Mark's team scores 25 2 pointers, meaning they scored $25 \times 2 = 50$ points in 2 pointers. His team also scores 8 3 pointers, meaning they scored $8 \times 3 = 24$ points in 3 pointers They scored 10 free throws, and free throws count as one point so they scored $10 \times 1 = 10$ points in free throws. All together his team scored $50 + 24 + 10 = 84$ points Mark's opponents scored double his team's number of 2 pointers, meaning they scored $50 \times 2 = 100$ points in 2 pointers. His opponents scored half his team's number of 3 pointers, meaning they scored $24 \div 2 = 12$ points in 3 pointers. They also scored half Mark's team's points in free throws, meaning they scored $10 \div 2 = 5$ points in free throws. All together Mark's opponents scored $100 + 12 + 5 = 117$ points The total score for the game is both team's scores added together, so it is $84 + 117 = 201$ points The answer is 201

Q: Bella has two times as many marbles as frisbees. She also has 20 more frisbees than deck cards. If she buys $\frac{2}{5}$ times more of each item, what would be the total number of the items she will have if she currently has 60 marbles?

A: Let's simplify the problem statement

without skipping any important information Bella has 60 marbles. She has twice as many marbles as frisbees. She also has 20 more frisbees than deck cards. If Bella buys $\frac{2}{5}$ times more of each item, how many marbles, frisbees, and deck cards will she have in total? Let's think step by step When Bella buys $\frac{2}{5}$ times more marbles, she'll have increased the number of marbles by $\frac{2}{5} \times 60 = 24$ The total number of marbles she'll have is $60 + 24 = 84$ If Bella currently has 60 marbles, and she has two times as many marbles as frisbees, she has $60 / 2 = 30$ frisbees. If Bella buys $\frac{2}{5}$ times more frisbees, she'll have $\frac{2}{5} \times 30 = 12$ more frisbees. The total number of frisbees she'll have will increase to $30 + 12 = 42$ Bella also has 20 more frisbees than deck cards, meaning she has $30 - 20 = 10$ deck cards If she buys $\frac{2}{5}$ times more deck cards, she'll have $\frac{2}{5} \times 10 = 4$ more deck cards. The total number of deck cards she'll have is $10 + 4 = 14$ Together, Bella will have a total of $14 + 42 + 84 = 140$ items The answer is 140

Question: A group of 4 fruit baskets contains 9 apples, 15 oranges, and 14 bananas in the first three baskets and 2 less of each fruit in the fourth basket. How many fruits are there?

A: Let's simplify the problem statement without skipping any important information There are 4 fruit baskets. The first three baskets contain 9 apples, 15 oranges, and 14 bananas each. The fourth basket has 2 less of each fruit than the first three. Find the total number of fruits. Let's think step by step For the first three baskets, the number of apples and oranges in one basket is $9 + 15 = 24$ In total, together with bananas, the number of fruits in one basket is $24 + 14 = 38$ for the first three baskets. Since there are three baskets each having 38 fruits, there are $3 \times 38 = 114$ fruits in the first three baskets. The number of apples in the fourth basket is $9 - 2 = 7$ There are also $15 - 2 = 13$ oranges in the fourth basket The combined number of oranges and apples in

the fourth basket is $7 + 13 = 20$ The fourth basket also contains $14 - 2 = 12$ bananas. In total, the fourth basket has $20 + 12 = 32$ fruits. The four baskets together have $114 + 32 = 146$ fruits. The answer is 146