

Q: Tim grows 5 trees. Each year he collects 6 lemons from each tree. How many lemons does he get in a decade?

A: The answer (arabic numerals) is

(Output) 3000 ✖

(a) Zero-shot

Q: Tim grows 5 trees. Each year he collects 6 lemons from each tree. How many lemons does he get in a decade?

A: **Let's give a correct and a wrong answer.**

(Output) Correct Answer:

Tim grows 5 trees and each tree gives him 6 lemons per year. So, in one year, he gets $5 * 6 = 30$ lemons. A decade is 10 years, so in a decade, he would get $30 * 10 = 300$ lemons. ✔

Incorrect Answer:

In one year, Tim gets $5 * 6 = 30$ lemons. A decade is 20 years, so in a decade, he would get $30 * 20 = 600$ lemons. (This is incorrect because a decade is 10 years, not 20 years.)

(b) Zero-shot-CP (Ours)

Figure 1: Example inputs and outputs of GPT-4 with (a) standard Zero-shot, and (b) ours (Zero-shot-CP). In contrast to Few-shot-CoT, which requires step-by-step reasoning examples for each task, our approach does not rely on any examples. Instead, we use the same prompt "Let's give a correct and a wrong answer" for all tasks, including arithmetic, symbolic, commonsense, and other logical reasoning tasks.