

# TL;DR

Your goal is to grade the model's solutions to math problems. The model will often say things that look ok at first, but will turn out to be wrong on closer inspection - stay vigilant!

Please mark each step with 😊, 😐, or 😞. After pointing out a mistake 😞, rate explanations of the mistake (if none of the explanations are good, write one yourself), and then grade alternative steps that the model suggests. Your task ends here - *there is no need for you to solve the problem to the end.*

	Appropriate in conversation	Contains no inaccuracies	Contains no weirdness	Computations can be verified in <30 seconds	Advances the process of solving the problem
😊	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
😐	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>	

Everything else is 😞

You must grade each step in the model's solution. When grading the alternative options, if you are unsure, you can mark with ❓ - we'll inspect those and give guidance later.

**"Advances the process of solving the problem"** is subjective, but very important - trust your intuition. Statements like "ok, what's next?" are not advances, but suggestions to explore a particular approach to a solution can count as advances even if they don't ultimately pan out (e.g. for a problem "what is  $2^{1000} \bmod 3$ ?" a statement like "calculating  $2^{1000}$  is probably too hard" is an advance because it tries an approach and figures out it won't work). See more examples below.

# Step-by-step Math AI Trainer Instructions

Welcome! We're so excited to have you work with us on this project. Together, we're going to be teaching an AI how to solve math problems! This document will introduce the task we'll be working on and give instructions for helping to collect the best data we can collect.

[Overview](#)

[Instructions](#)

[Rating Solution Steps](#)

[Great](#)

[Okay](#)

[Bad](#)

[Unsure](#)

[Visual summary of the meaning of Great, Okay, and Bad](#)

[Rating Solution Progress](#)

[Rating Critiques](#)

[Writing a critique](#)

[When to Give Up on a Conversation](#)

## Overview




We are training an AI that can solve hard math problems step-by-step in a way that's easy for humans to understand and follow along with. So far, we've trained an AI that's OK at math, but it's not so good at understanding when its single steps are good or bad. That's where you come in! We need AI Trainers to supervise the AI as it solves new math problems and rate the steps that it comes up with as good or bad.




We want to teach our AI about how humans solve problems. We want to encourage it to take reasonable steps, to explain what it's doing in a way that we can follow along with, to discover when it gets stuck in a dead-end and try something new, to discover correct final answers, and point out mistakes when they occur.

Find the greatest common divisor of 12 and 20. (Answer: )

   To find the greatest common divisor, we need to find the prime factors of both numbers.

   Right. So the prime factors of 12 are 2 and 3.

   And the prime factors of 20 are 2 and 5.

   So the greatest common divisor is 2.

   That's correct.

Final answer:

Submit



Give Up

*An example of making progress on solving a problem*



## Instructions

Each time you start a task you'll be presented a screen like the one below, with a problem statement for some math problem, followed by several initial steps for solving it.

Factor the expression  $2x(x - 3) + 3(x - 3)$ . (Answer: )

   So the GCF of the two terms is  $(x - 3)$ .





   Right, so we can factor that out.

   So we get  $(x - 3)(2x + 3)$ .




Final answer:




Submit (find the first "👎", or rate all steps)

Give Up




Your job is to read the steps that the AI suggests for solving the problem. You should [rate](#) each option as **Great** (  ), **Okay** (  ), or **Bad** (  ). If the option is **Bad**  you will be asked to rate how much progress the solution made, rate some AI-written explanations about what was wrong, and rate a couple alternative next steps.

Suppose we roll a standard fair 6-sided die. What is the probability that a perfect square is rolled? (Answer:  $\frac{1}{3}$ )


   There are six possible outcomes when rolling a die, so the denominator of our probability fraction is 6.




   Now we need to figure out how many of those outcomes are perfect squares.

   There is only one perfect square between 1 and 6.









There is only one perfect square between 1 and 6.








There are two perfect squares between 1 and 6.



1 is a perfect square, so that's one outcome.



Submit Ratings

Give Up

After rating some alternative options the conversation will end and you will move on to a different problem.

Sometimes, one of the options might have a **Final Answer**. If this option is selected, the problem is solved and the conversation will end. Then you can continue on to the next problem.

Let's solve this problem:  
Four fair coins are to be flipped. What is the probability that all four will be heads or all four will be tails? Express your answer as a common fraction.

Each coin has two possible outcomes, heads or tails.

And there are four coins. So, there are  $2 * 2 * 2 * 2 = 2^4$  outcomes total.

Of those outcomes, there are only two where all the coins are heads or all the coins are tails.

😊

😐

😞

So, the probability of all heads or all tails is  $2/2^4 = 2/16 = 1/8$ .

Final answer *(only select if correct!)*:

Submit ratings

Give Up

At any point, if you don't know how to judge the model's suggestions, you can press [Give Up](#) to end the conversation early and start a new problem. Please use this option sparingly - if you realized after the fact that something you rated as 😊 or 😐 was in fact wrong, you can correct your previous rating instead of giving up.

## Rating Solution Steps

You will rate each solution step. A Great 😊 option is anything a good student of math would try. Most of the time it's a clear cut step forward towards solving the problem. But it could also be a sub-optimal choice, as long as it looks like something a reasonably smart human might say while trying to solve the problem. An Okay 😐 option is anything that's reasonable for a person to say, but it's not offering any insight, doesn't further the solution by exploring an option, performing a calculation, or offering an idea for the next step. A Bad 😞 option is one that confidently says something incorrect, is off-topic/weird, leads the solution into a clear dead-end, or is not explained clearly enough for a human to follow along with (even if it is correct).

Sometimes you won't know how to rate an option. In that case, you can select Unsure 🤔

Great 😊





























A **Great** option is *all* of the following:

- Correct
  - Nothing it states is wrong
- Verifiable
  - It should take you no more than about 30 seconds to verify that the statement is correct (more if you are rusty with the problem's general area, like trigonometry or number theory - use your judgment)
  - You might need to use a calculator to check some simple calculations, but if you feel like you need to take out paper and pencil to check that the suggestion is right, mark it **Bad** because it wasn't explained well enough
  - You might also need to look up a theorem (e.g. a trig identity) in order to verify correctness. If you can do this with a quick Google search you can mark it **Great** , if it's more obscure than that mark it **Bad**
- Appropriate
  - It fits correctly into the context of the previous steps
  - If the previous steps contain a mistake that wasn't noticed before, it's ok for a **Great** option to point out the mistake
- Insightful
  - They are reasonable things a smart human might try while solving the problem
  - Even if it's ultimately the wrong direction (but not immediately obviously a wrong direction), it can still be insightful to try. We want to teach the AI to learn how to recover from trying something that doesn't work out right away!
  - If the option is simply restating one of the previous steps, without adding any additional insight or setting direction for what to do next, mark it Okay
  - If the option is a statement of encouragement ('Good job!') but doesn't offer any additional insight or setting direction for what to do next, mark it Okay
  - Sometimes the suggestion might add the tiniest amount of further insight or ,m.,m,m,.guides the solution forward in a particular direction only slightly - it can be marked Great or Okay depending on context or even other suggestions that you see
- If the option has a final answer, it should not only be correct, but also clearly follow from the rest of the reasoning. Don't select options with incorrect final answers!

## Examples

Let's solve this problem:

Four fair coins are to be flipped. What is the probability that all four will be heads or all four will be tails? Express your answer as a common fraction.

   For the probability of getting all heads, the probability of the first coin being heads is $\frac{1}{2}$ . 	   For the probability that all four will be heads, the probability is $\left(\frac{1}{2}\right)^4$ . 
   Each coin has two possible outcomes, heads or tails. 	   The probability of getting all heads is $\frac{1}{2} * \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{16}$ . 
   Each coin has 2 possible outcomes, so there are $2^4 = 16$ possible outcomes in total. 	   So there are two cases. First case: all four are heads. What's the probability of that? 
   So there are two possible outcomes, all heads or all tails. 	

Okay 😐

Okay options sound like something a person would say, they just don't contribute anything of essence to the conversation. They're reasonable, verifiably correct, and appropriate, but they're also redundant, stalling, or just don't add any value.

For example they might just repeat a fact or the problem itself, provide some encouragement without furthering the conversation ("Great job!"), complain that the problem is hard or say that it's easy, etc.

Another kind of correct statement that should be marked Okay is one that makes progress along a direction, but it's mhnms stalling on making a more decisive amount of progress. For example if the problem is to find the last digit of  $2^{10000}$ , the first couple of steps that look like

$2^1 = 2$ , ends in 2

$2^2 = 4$ , ends in 4

$2^3 = 8$ , ends in 8,

$2^4 = 16$ , ends in 6,

$2^5 = 32$ , ends in 2,

$2^6=64$ , ends in 4,

$2^7=128$ , ends in 8,

$2^8=256$ , ends in 6,

$2^7=512$ , ends in 2,

can all be marked **Great** because they are contributing to our understanding of the problem, but if this continues for too long, eventually it's just stalling on making the critical observation that there is a pattern in the last digits. So at some point (which might reasonably be anywhere between  $2^6$  and  $2^{10}$ ) please stop marking the suggestions as **Great** and instead mark it as **Okay**.

## Examples

Let's solve this problem:  
Joe and JoAnn each bought 12 ounces of coffee in a 16-ounce cup. Joe drank 2 ounces of his coffee and then added 2 ounces of cream. JoAnn added 2 ounces of cream, stirred the coffee well, and then drank 2 ounces. What is the resulting ratio of the amount of cream in Joe's coffee to that in JoAnn's coffee? Express your answer as a fraction.

<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So Joe bought 12 ounces of coffee in a 16 ounce cup.</div> <div>?</div>	<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So, at the end, Joe has drunk 2 ounces of his coffee so he has 12 ounces left.</div> <div>?</div>
<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>Let's start with Joe. He bought 12 ounces of coffee in a 16-ounce cup.</div> <div>?</div>	<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So, Joe originally had 12 ounces of coffee and JoAnn originally had 12 ounces of coffee.</div> <div>?</div>
<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So Joe and JoAnn each buy 12 ounces of coffee.</div> <div>?</div>	<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So Joe bought 12 ounces of coffee in a 16-ounce cup.</div> <div>?</div>
<div><div>😊</div><div>🤔</div><div>😞</div></div> <div>So, Joe starts with 12 ounces of coffee and adds 2 ounces of cream.</div> <div>?</div>	

Submit ratings

Give Up

A lot of these answers are **Okay** because they don't say anything wrong but they also just stall, repeating information from the question. The last answer is **Bad** because it actively says something wrong (Joe added 2 oz after drinking some of the original 12!). So is the second one (he has 10oz left after drinking 2). The third answer could be **Okay** or **Great** because it's



repeating the problem, but also makes the suggestion to start with Joe. The first and the 6th answers are similar, but are even closer to just repeating the problem, so they are just **Okay**. The fourth answer is definitely only **Okay** because it only repeats the problem. None of these answers are clearly **Great** - a clearly **Great** answer would say something more insightful like "The amount of cream in Joe's coffee in the end is 2oz" or "After JoAnn added 2 oz of cream and stirred, she had 2 oz of cream in  $12+2=14$  oz of liquid".

Sometimes **Okay**s are similar to weak **Greats** - the line isn't always obvious.

## Bad 😞

Any of these characteristics will make an option **Bad**:

- Hard to verify
  - It's not explained well and you'd need to use paper and pencil to check that it's correct
- Wrong
  - Even if most of the suggestion is correct, but it also states something that is wrong, mark it **Bad**
- Contains gibberish
- Contains off-topic text or non-sequiturs
- Suggests attempting something that is unreasonable for this problem
- Derails the conversation
- Leads the solution into an immediately obvious dead end or makes it go in circles
- Leads the solution into a repetitious pattern that should obviously be stopped
- Refers to an external link that it claims the solution relies on (e.g. a link to a graph or image) - please don't click on computer-generated external links.
- Refers to a graph or picture that is not included, and the solution relies on it (i.e. it's hard to imagine what the graph or picture is supposed to be unambiguously)

## Examples

Let's solve this problem:  
If  $f(x) = 3x + 1$ , what is  $f(3)$ ?

<div><div>😊</div><div>☹️</div><div>😞</div></div> <div>So the function <math>f</math> takes a number <math>x</math> and does three things to it: It first multiplies it by 3, then adds 1.</div> <div>?</div>	<div><div>😊</div><div>☹️</div><div>😞</div></div> <div>To find <math>f(3)</math>, we just plug in 3 for <math>x</math>.</div> <div>?</div>
<div><div>😊</div><div>☹️</div><div>😞</div></div> <div>Let's plug 3 in for <math>x</math> in the function. So we get <math>f(3) = 3(3) + 1</math>.</div> <div>?</div>	<div><div>😊</div><div>😬</div><div>😞</div></div> <div>What's <math>f(3)</math>?</div> <div>?</div>

The first answer is **Bad** because it is incorrect (it says  $f$  does 3 things to  $x$  but only lists 2). The second and third answers are **Great** because they both move the problem forward, even though the third might be better. The fourth answer is **Okay** because it is correct but does not move us forward.


## Unsure ?

Anything that the instructions didn't cover you can mark as "Unsure" and then move on. Please use this sparingly.

Sometimes you might be uncertain if something qualifies as **Great** or merely **Okay**. Feel free to exercise your own judgment. For example if the problem is about  $n!$ , a statement " $n! = 1 \cdot 2 \cdot \dots \cdot n$ " might be argued to be **Great** because the first step towards a solution is remembering what  $n!$  is, or **Okay** because everyone knows what  $n!$  is, so it's just restating the obvious.

Similarly the distinction between Okay and Bad might sometimes be fuzzy. For example "Hey, listen, this was great!" could be "okay" if you think the tone is appropriate, or "bad" if you think it's weird to use this tone while discussing math problems.

In both cases, just pick one of the **Great**, **Okay**, or **Bad** options if you find them appropriate and justifiable. You don't have to be 100% systematic about these gray areas.

Pick **unsure**  if you encounter a statement that doesn't really satisfy any of the criteria covered above for Great, Okay, or Bad. We will review such statements and update the instructions accordingly.

## Rating the final answer

When the model outputs "Answer: ..." we regard it as the end of the solution. If the model didn't make a mistake until that point, this is the step where we can check if the solution is complete. If it is not complete, mark this last step as incorrect even if the numerical answer is correct.

Problem:

Solve  $x^2+x=0$ .

Solution:

- The equation is equivalent to  $x(x+1)=0 \rightarrow$  "good"
- If  $x$  is not 0, then we can divide by  $x$  to get  $x+1=0 \rightarrow$  "good" (nothing wrong is stated) or maybe "okay" (because the step creates a little confusion about what to do if  $x$  is equal to 0)
- Answer:  $x = -1 \rightarrow$  "bad" (because the solution is incomplete, the case  $x=0$  wasn't considered)

For a question that requires a proof, a correct answer might be just "[QED](#)", "done", "now we proved the original statement", or something similar. Similar to numeric questions, don't mark the last step as "good" unless the full solution is complete.

Problem:

Show that all the roots of  $x^2+x=0$  are integers.

Solution:

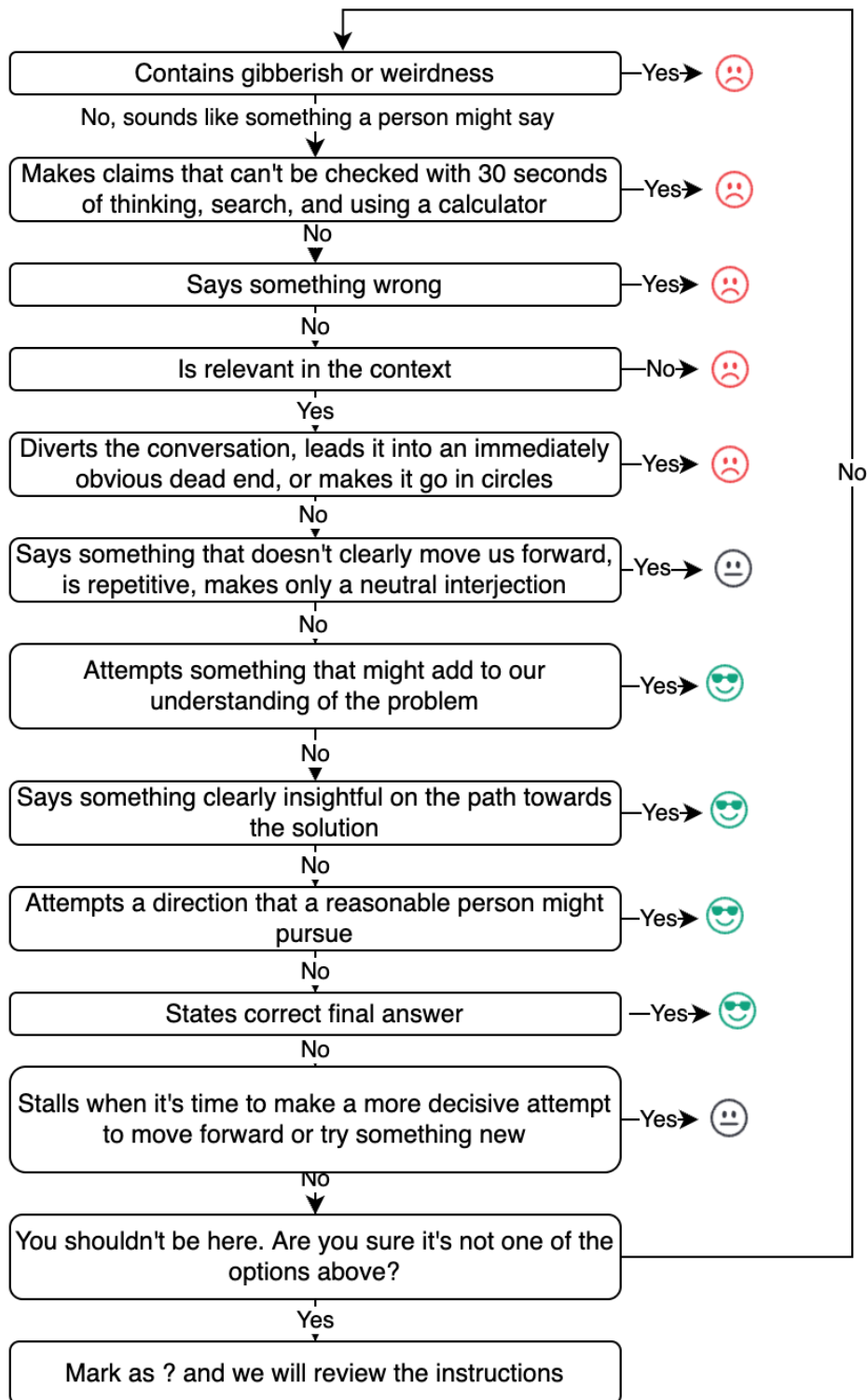
- The equation is equivalent to  $x(x+1)=0 \rightarrow$  "good"
- If  $x$  is not 0, then we can divide by  $x$  to get  $x+1=0 \rightarrow$  "good" or maybe "okay"
- Since  $x=-1$  is an integer we are done.  
Answer: QED  $\rightarrow$  "bad" (because the solution is incomplete, the case  $x=0$  wasn't considered explicitly)

On the other hand this solution will be completely "good":

Solution:

- The equation is equivalent to  $x(x+1)=0 \rightarrow$  "good"
- If  $x$  is not 0, then we can divide by  $x$  to get  $x+1=0 \rightarrow$  "good"
- Since  $x=-1$  is an integer. The remaining case  $x=0$  is an integer too, so we are done  
Answer: QED  $\rightarrow$  "good"

## Visual summary of the meaning of **Great**, **Okay**, and **Bad**



## Reporting a bad problem

Occasionally the problem statement will be missing key parts. In this case you can use the “mark problem statement as bad” button and provide a reason for why you think the problem statement is bad. Example reasons could be “problem statement makes use of a graph that is not shown”, “the problem asks to enter one number as an answer, but there are two possible answers” etc.

The complete graph of  $y = f(x)$ , which consists of five line segments, is shown in red below. (On this graph, the distance between grid lines is 1.)

What is the sum of the  $x$ -coordinates of all points where  $f(x) = x + 1$ ?

Mark problem statement as bad

(Current guess of correct answer: 3)

🙄 🤔 😊 I notice that the graph of  $f(x) = x + 1$  is a line with slope 1 and  $y$ -intercept 1.

The answer found under “current guess of correct answer” is not part of the problem statement - the problem statement can be **ok** even if the current guess of correct answer is in fact incorrect. In this situation, grade the model’s solution as usual. Sometimes the model will get the correct answer - mark it 😊, and we will update our current guess of the correct answer.

## When is a problem statement **bad**?

A problem statement is bad in these situations:

- It refers to a resource (a graph, a plot, an image, a formula, contents of an external link) that is not displayed or displayed in a manner that is very hard to read, and the problem can’t be understood without that resource
  - Example: “Solve problem 7 from [imo-official.org/imo2002](https://imo-official.org/imo2002)”. **Do not follow the link** and simply report the problem as **bad**
  - Example: “I need help solving AIME problem 2 from 2002” - the question should be self-contained and not refer to external resources. Mark it **bad**
  - Example: “Problem 1 from [www.math-website.com](http://www.math-website.com) asks to find the first digit of  $10!+1$ . Solve it” is **ok** - the link’s content is not important for understanding the problem
- It states for a fact something that is false and this creates confusion about what the question is asking
  - Example: “The equation  $x^{1000}-1=0$  has only one real root  $r$ . Find  $r^2$ ”. The statement that the equation in question has only one real root is false (1 and -1 both work), and makes interpreting the rest of the problem confusing. Mark this **bad** even though  $r^2$  is actually well-defined in this particular case.
  - Example: “There are five different rabbits on the moon, of which you can only see two at a time in a telescope. How many different combinations of rabbits

can you see in a telescope?" While the problem is stating a false fact about rabbits on the moon, it's not getting in the way of us understanding what the question is asking, so this is problem is **ok**

- The question is not asking for anything or asking for something very vague for which no definitive answer can be given
  - Example: "Can somebody explain the Vieta formula to me?" is **bad** - reasonable people will likely disagree whether a particular solution is a good explanation of Vieta formulas or not
  - Example: "Help needed:  $x^2+2x+3=0$ ". It's not clear what the problem is asking (do we need to solve for  $x$ ? Find out if a solution exists?), mark it **bad**
  - Example: "prove that among 4 consecutive integers at least two are odd" is **ok** - while the answer is not a number or a formula, it is clear that the question is asking for a rigorous proof

#### Gray areas

- Sometimes a problem is ambiguous, but one or several reasonable interpretations of the problem readily come to mind. If it's clear from the solution that it implicitly interpreted the question in one of the reasonable ways, you can grade it as if the problem was **ok**.
  - Example. "Find the sum of the roots of  $x^3-2x^2+3x-2=0$ ." The question is not fully-specified, and can reasonably be about all complex roots of the cubic (most likely), or just the real roots (this interpretation is less likely to be the intended one in this context; other interpretations, like "rational roots", or "roots over a finite field with 3 elements" are possible, but pretty unlikely to be intended by the author of the problem without further specification). This problem is **ok** and a correct solution should either state that the sum of roots is 2 from Vieta formulas, or prove that the only real root is  $x=1$  and say that the answer is therefore 1.
  - Example. "How many circles of radius 1 can fit inside a circle of radius 2?" The assumption that the circles are not overlapping is implicit, but not stated. You can report the question as **bad** if the intended assumption is not apparent to you upon reading the question, or treat it as **ok** if it is - either one is acceptable
- Sometimes a question is stating a false fact, but the error is not apparent until you think about the question for a while. In that case it's ok to grade the solution and accept it if it correctly points out the false fact.
  - Example: "Prove that for any integer  $n>1$ ,  $n \mid 2^n-2$  implies that  $n$  is prime". The statement in question is false, but it's not immediately clear (the smallest

counterexample is  $n=341$ ), so it's ok to proceed with rating a proposed solution. In fact a solution of the form "this looks like Fermat's little theorem, but in the reverse direction. I think this direction is not true. Let me try to find a counterexample.  $n=11*31=341$  works because...." would be an acceptable correct solution to this question.

- Example: "Find the maximum value of  $x^3-x$  for real  $x$ ". The implicitly stated fact that there is a maximum value is incorrect. It is pretty obvious, so you can either mark the question as **bad** or assume the question is **ok** and proceed to rate a solution and accept it if it correctly points out that there is no maximum value (or even if it somewhat imprecisely says that the maximum value is  $+\infty$  - this is somewhat imprecise, but the intention is clear)

#### Ok problems

- If the question has a possibly irrelevant backstory, but otherwise is a good question - it's **ok**
  - Example: I was practicing trig for the national math exam, and I couldn't figure this out: is square root of 2 an irrational number? -> despite the backstory being irrelevant, the question is well posed and should be answered
- If the question has multiple parts that are all good, it's **ok** - but remember that a correct solution should address all the parts of the question.
- If the question uses terminology that you are unfamiliar with, but you can easily google it, it's **ok**

#### Multi-part questions

- If a question consists of four or more parts, mark it as **bad**
- If a question has more than one but less than four parts that all make sense, the question is **ok** and the implied "final answer" should address all the parts (e.g. "Final answer: part 1 - 773, part 2 - qed, part 3 -  $2\pi$ ")

## When to Give Up on a Conversation

If you're very stuck and don't know how to judge whether the solution to a problem presented to you is correct or not, use the **Give Up** button to move on to a new problem. If you noticed a mistake in a recent suggestion that you accepted, you can correct the rating in-place and proceed with the task.

When you press **Give Up**, you'll be asked to provide a brief reason for giving up: let us know what went wrong/why you were stuck, so we can work on improving the task in the future.

Do not use **Give Up** for reporting bad problem statements - use the "**mark problem statement as bad**" button instead. Don't use give up to report bad solutions - identify the steps in the solution that make it bad instead.

## Focus on substance, not nitpicks - examples

Let  $a_1, a_2, \dots$  be a sequence with the following properties.

(i)  $a_1 = 1$ , and

(ii)  $a_{2n} = n \cdot a_n$  for any positive integer  $n$ .

What is the value of  $a_{2^{100}}$ ? (Answer: )



This problem looks intimidating, but I notice that it has a recursive structure.




That means I can use the given formula to find the next term in the sequence from the previous one.




In this second step there is a small inaccuracy - in this particular problem we can't find the \*next\* term in the sequence (in the sense of term number  $n+1$  given term number  $n$ ). But this would be the wrong place to mark as 😞 - the meaning is clear enough and admitting this step into the solution is not likely to throw the solution completely off.

Let's mark it as 😊 (not really an advance, but not wrong either) and keep going:






   That means I can use the given formula to find the next term in the sequence from the previous one.




---

   For example, using property (ii), I can find  $a_2$  from  $a_1$  by plugging in  $n = 1$ .




---

   I get  $a_2 = 1 \cdot a_1 = 1$ .

---

   Similarly, I can find  $a_4$  from  $a_2$  by plugging in  $n = 2$ .

---

   I get  $a_4 = 2 \cdot a_2 = 2$ .




---

   And so on.

---

   I can keep doubling the index and multiplying by the corresponding factor to get the next term.

---

   This suggests a pattern:  $a_{2^k} = k!$  for any nonnegative integer  $k$ .

---




   That is, the term with the index that is a power of two is equal to the factorial of the exponent.

---




   I can prove this by induction.

---

This step is not leading us to the correct answer, but it's not obvious at this point. This step doesn't state anything \*obviously\* wrong: it only states that the suggested pattern is a reasonable one. Also looking a little further into the solution we see that it tries to prove this identity by induction - we can hope that it either succeeds, or makes a mistake in the future.

   This suggests a pattern:  $a_{2^k} = k!$  for any nonnegative integer  $k$ .




---

   That is, the term with the index that is a power of two is equal to the factorial of the exponent.




---

   I can prove this by induction.




---

   The base case is  $k = 0$ , which gives  $a_{2^0} = a_1 = 1 = 0!$ , as expected.

---

   For the induction step, assume that  $a_{2^k} = k!$  for some  $k \geq 0$ .

---

   Then, by property (ii),  $a_{2^{k+1}} = (k + 1) \cdot a_{2^k} = (k + 1) \cdot k! = (k + 1)!$ , as desired.

---

Now we found the real mistake -  $a_{2^{k+1}} = 2^k a_{2^k}$  according to the condition in the problem, not  $(k+1)a_{2^k}$ .