Step-by-Step Verification Guidelines

Tracking shuffled objects

Overview

We've used an AI model to generate a sequence of steps to complete a task, but some of these steps might be wrong. Your task is to **go through each step and verify whether it is correct or not.**

The AI's task is to track the movements of a set of objects or people. The given context always contains a sequence of changes, and the steps should track each change sequentially. For example:

Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a yellow ball, Bob has a blue ball, and Claire has a pink ball.
As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Alice and Bob swap balls. Finally, Claire and Bob swap balls. At the end of the game, Bob has the
Options:
(A) yellow ball
(B) blue ball
(C) pink ball

At the start: Alice: yellow, Bob: blue, Claire: pink.	☺ 😐 ☹
Claire and Alice swap balls: Alice: pink, Bob: blue, Claire: yellow.	☺ 😐 ☹
Alice and Bob swap balls: Alice: blue, Bob: pink, Claire: yellow.	☺ 😐 ☹
Claire and Bob swap balls: Alice: blue, Bob: yellow, Claire: pink.	☺ 😐 ☹
At the end of the game, Bob has the yellow ball. So the answer is (A)	☺ 😐 ☹

Instructions

(1) You should label each step as positive or negative.
Once you label one step, the next step will be unlocked.

Positive	Negative
 This step is logically correct and contributes to task progression. Typos and grammatical mistakes are allowed as long as the logic is still correct.	 This step is **one or more** of the following: ● Logically incorrect or requiring large logical jumps ● Repetitive or redundant ● Off-topic or contains gibberish ● Leads to an obvious dead end ● Refers to external links, images, or graphs

(2) Each set of steps can be submitted when **either**:
   (a) All the steps have been labelled as positive; **or**
   (b) The last labelled step is negative. In this case, the next step will not unlock and you do not have to label the remaining steps.

Example labels

In this section, we've taken the "positive" steps for this example and added some possible variations, as well as some "negative" examples so you can compare them side-by-side. Note that you will not be able to do this when doing the annotation, so make sure to check each step carefully!

Alice, Bob, and Claire are playing a game. At the start of the game, they are each holding a ball: Alice has a yellow ball, Bob has a blue ball, and Claire has a pink ball.
As the game progresses, pairs of players trade balls. First, Claire and Alice swap balls. Then, Alice and Bob swap balls. Finally, Claire and Bob swap balls. At the end of the game, Bob has the
Options:
(A) yellow ball
(B) blue ball
(C) pink ball

		**Reason**

At the start: Alice: yellow, Bob: blue, Claire: pink.	At the start: **A:** yellow, **B:** blue, **C:** pink.	It is fine to use shorthand for names and/or objects, as long as the association is clear (e.g. A is for Alice rather than Bob) and the logic is correct.
Claire and Alice swap balls: Alice: pink, Bob: blue, Claire: yellow.	Claire and Alice **swaps** balls: Alice: pink, Bob: blue, Claire: yellow.	Grammatical mistakes are fine, as long as the sentence is easily understandable and the logic is correct.
Alice and Bob swap balls: Alice: blue, Bob: pink, Claire: yellow.	Alice and Bob swap balls. **[MISSING]**	Steps can be split up into smaller steps, as long as there are no jumps in logic. This example is considered positive since the logic is correct and there are no jumps (so far).
Claire and Bob swap balls: Alice: blue, Bob: yellow, Claire: pink.		
At the end of the game, Bob has the yellow ball. So the answer is (A)		

☺	☹	**Reason**
At the start: Alice: yellow, Bob: blue, Claire: pink.		
Claire and Alice swap balls: Alice: pink, Bob: blue, Claire: yellow.	Claire and Alice swap balls: **Alice: blue, Bob: yellow**, Claire: pink.	The objects were swapped incorrectly.
Alice and Bob swap balls: Alice: blue, Bob: pink, Claire: yellow.	**Claire and Bob swap balls**: Alice: blue, Bob: pink, Alice: blue.	The statement does not match the swap described in the question.
Claire and Bob swap balls: Alice: blue, Bob: yellow, Claire: pink.	**At the end of the game, Bob has the yellow ball. So the answer is (B)**	There is a logical jump between this step and the previous step, because the swap between Claire and Bob was ignored.

		Even though the answer is correct, this step is incorrect.
At the end of the game, Bob has the yellow ball. So the answer is (A)	At the end of the game, Bob has the yellow ball. So the answer is **(B)**	The statement “Bob has the yellow ball” is correct, but the chosen option is wrong.
	Final answer: Bob has the yellow ball.	Steps that are redundant (e.g. repeating previous steps) are considered negative.
	Bob has the yellow ball. ABCDEFG	Any unrelated text is considered negative.