

## A Supplementary

### A.1 Comprehensive Comparison of Length, Rewards, and Skip Ratio Across Three Benchmarks

We present a comprehensive comparison of length, rewards, and skip ratio across three benchmarks. The results reveal a consistent trend: TON reduces completion length and increases the skip ratio as rewards increase during training.

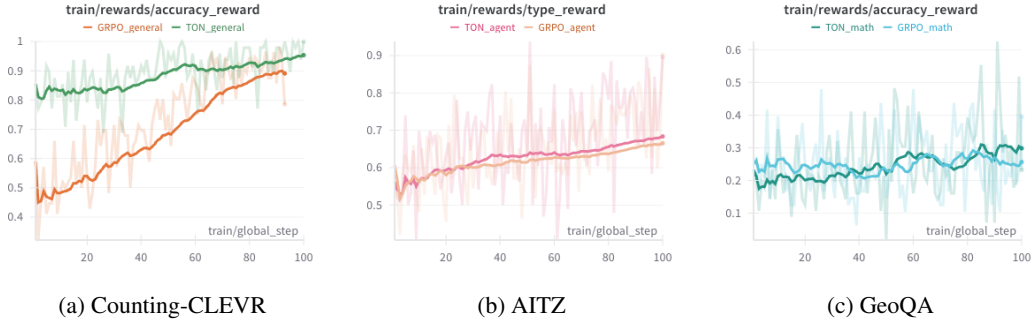


Figure 7: **Rewards of the output** during our TON training on three benchmarks.



Figure 8: **Completion length of the output** during our TON training on three benchmarks.

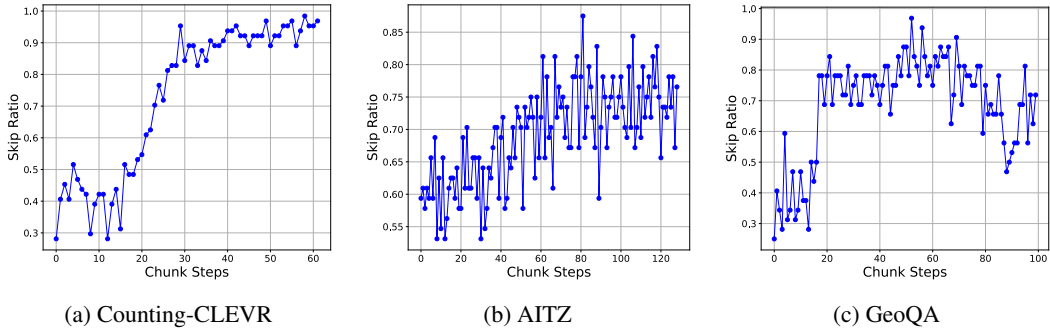


Figure 9: **Skip ratio of the output thinking** during our TON training on three benchmarks.

### A.2 Reward for length rather than SFT

We give the reward  $r_l = 1$  for the model if it outputs `<think>\n\n</think>`, and otherwise 0. We observe that the length reward remains at 0 during the first 100 steps. The visualization of the entire training process is shown below, highlighting our proposed thought dropout in the SFT stage.



Figure 10: Length penalty rewards and completion length on AITZ datasets on Qwen2.5-VL-3B.

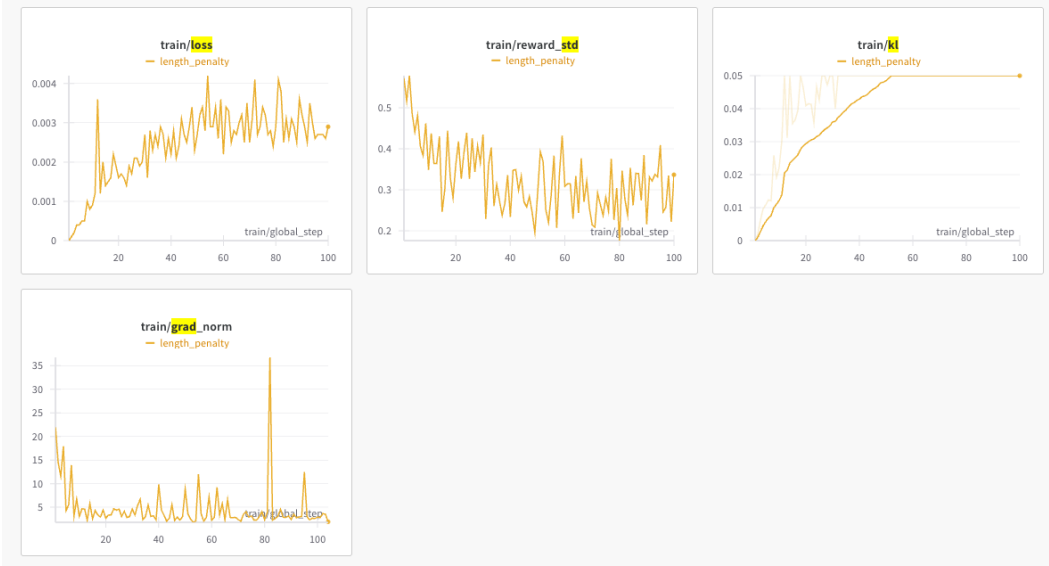


Figure 11: Length penalty rewards and completion length on AITZ datasets on Qwen2.5-VL-3B.

### A.3 More cases

We give more cases to show the effectiveness and efficiency of our proposed TON.



Figure 12: TON applied to the AITZ out-of-distribution domain: the task is to uninstall the messaging apps. TON performs well without extensive reasoning.

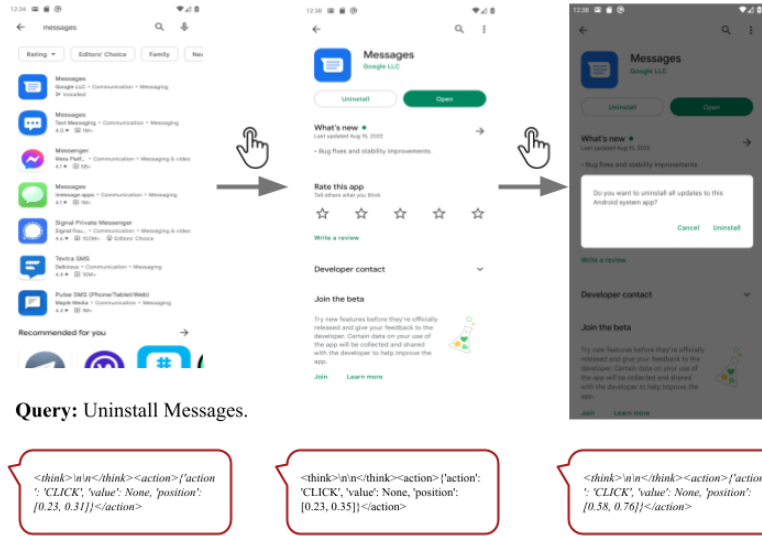
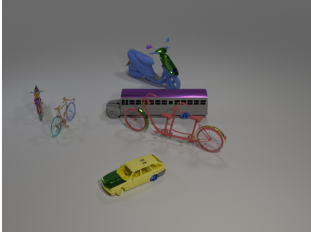
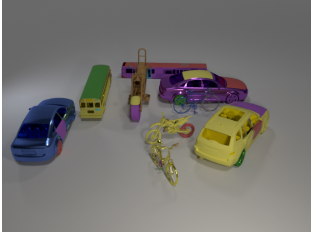
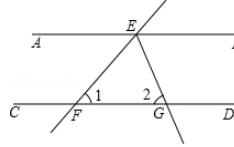


Figure 13: TON applied to the AITZ out-of-distribution domain: the task is to uninstall the messaging apps. TON performs well without extensive reasoning.

Table 5: Counting example from SuperCLEVR [15]. Although the question is out-of-distribution, the performance of TON and vanilla GRPO differs significantly in their outputs. The output from vanilla GRPO is excessively lengthy (over 500 tokens), focusing more on reasoning rather than providing a direct answer. In contrast, TON delivers a concise response, effectively bypassing the lengthy reasoning process.

Counting examples from SuperCLEVR [15]		
		
Question	How many objects are either small green wagons or buses?	How many other objects are there of the same color as the matte car?
GRPO's outputs	<p>"&lt;think&gt;, let's tackle this question step by step. The user is asking how many objects are either small green wagon-sör buses. I need to check each object in the image provided., let's parse the image. There are six objects listed:1. A large blue rubber scooter with specific coordinates and rotation.2. A small yellow rubber car (wagon) with details.3. A small brown rubber bicycle.4. A large red rubber tandem bicycle.5. A small gray metal dirt bike.6. A large gray metal school bus.question is about two categories: small green wagons and buses. Let's break down each part.with small green wagons: Looking at the list, there's a small yellow rubber car which is a wagon. The color here is yellow, not green. So that doesn't fit. Are there any other small green wagons? The other objects are all different colors and types. The only green object mentioned is the first one, but it's a scooter, not a wagon. So no small green wagons exist in the scene., buses: The sixth object is a large gray metal school bus. Since the question says "</p>	
TON's outputs	<think>\n\n</think> <answer>1</answer>	<think>\n\n</think> <answer>3</answer>

Q: As shown in the figure,  $AB \parallel CD$ , straight line  $EF$  intersects  $AB$  at point  $E$ , intersects  $CD$  at point  $F$ ,  $EG$  bisects angle  $BEF$ , and it intersects  $CD$  at point  $G$ , angle  $1 = 50.0$ , then angle  $2$  is equal to ()



TON:

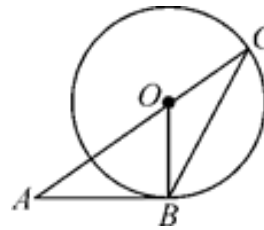
<think>\n\n</think><answer>65</answer>

GRPO:

<think>\nTo solve for angle 2, we need to use the properties of parallel lines and angle bisectors.\n\n1. Since  $AB$  is parallel to  $CD$ , and  $EF$  is a transversal, angle  $BEF$  and angle  $BFG$  are supplementary angles because they form a linear pair.\n2. Given that angle 1 is 50 degrees, and since  $EG$  bisects angle  $BEF$ , angle  $BEG$  must be half of angle  $BEF$ .\n3. Therefore, angle  $BEG = 90 - 50 = 40$  degrees.\n4. Since  $EG$  bisects angle  $BEF$ , angle  $BEG =$  angle  $FEG = 40$  degrees.\n5. Now, angle 2 is the same as angle  $FEG$  because they are corresponding angles formed by the transversal  $EF$  intersecting the parallel lines  $AB$  and  $CD$ .\n6. Therefore, angle 2 = 40 degrees.\n\n</think>\n\n<answer>\n40\n</answer>

Figure 14: Comparison between TON and vanilla GRPO in GeoQA

Q: As shown in the figure,  $AB$  is tangent to circle  $O$  at point  $B$ , and the extended line of  $AO$  intersects circle  $O$  at point  $C$ . Connect  $BC$ , if angle  $A = 36.0$ , then angle  $C$  is equal to ()



TON:

<think>\n\n</think><answer>27</answer>

GRPO:

<think>\nTo solve for the measure of angle  $C$ , we need to use properties of circles and tangents.\n\n1. \*\*Identify the given information:\*\*\n -  $AB$  is tangent to circle  $O$  at point  $B$ .\n - Angle  $A$  is 36 degrees.\n - We need to find the measure of angle  $C$ .\n2. \*\*Use the property of the tangent:\*\*\n - Since  $AB$  is tangent to the circle at point  $B$ , the radius  $OB$  is perpendicular to  $AB$ . Therefore, angle  $OBA$  is 90 degrees.\n3. \*\*Determine angle  $OAB$ :OAB, the sum of the angles is 180 degrees.\n - So, angle  $OAB$  + angle  $A$  + angle  $OBA$  = 180 degrees.\n - Substituting the known values: angle  $OAB$  + 36 degrees + 90 degrees = 180 degrees.\n - Simplifying, we get angle  $OAB = 54$  degrees.\n4. \*\*Identify angle  $COB$ :OC is a radius and  $OB$  is also a radius, triangle  $OBC$  is isosceles with  $OB = OC$ .\n - Therefore, angle  $OCB =$  angle  $OBC$ .\n5. \*\*Determine angle  $COB$ :COB subtends the same arc as the inscribed angle  $CBA$ .\n - The measure of the central angle is twice the measure of the inscribed angle that subtends the same arc.\n - So, angle  $COB = 2 *$  angle  $CBA$ .\n6. \*\*Relate angles in triangle  $OBC$ :OBC, the sum of the angles is 180 degrees.\n - So, angle  $OCB$  + angle  $OBC$  + angle  $COB = 180$  degrees.\n - Since angle  $OCB =$  angle  $OBC$ , we can write:  $2 *$  angle  $OCB$  + angle  $COB = 180$  degrees.\n - Substituting angle  $COB = 2 *$  angle  $CBA$ , we get:  $2 *$  angle  $OCB$  +  $2 *$  angle  $CBA = 180$  degrees.\n - Simplifying, we get:  $2 *$  angle  $OCB$  +  $2 * 36$  degrees = 180 degrees.\n - This simplifies to:  $2 *$  angle  $OCB$  + 72 degrees = 180 degrees.\n - Solving for angle  $OCB$ , we get:  $2 *$  angle  $OCB = 108$  degrees.\n - Therefore, angle  $OCB = 54$  degrees.\n\nSo, the measure of angle  $C$  is 54 degrees.\n\n</think>\n\n<answer>\n54\n</answer>

Figure 15: Comparison between TON and vanilla GRPO in GeoQA

Q: As shown in the figure, the diameter CD of circle O crosses the midpoint G of chord EF, angle DCF = 20.0, then angle EOD is equal to ()



TON:

<think>\n\n</think><answer>40</answer>

GRPO:

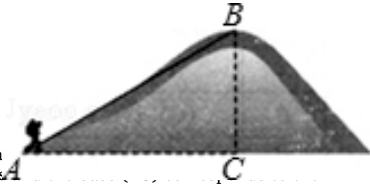
<think>\nTo solve for the measure of angle EOD, we need to use properties of circles and angles.\n\n1. \*\*Identify the given information:\*\*\n -  $CD$  is the diameter of the circle.\n -  $G$  is the midpoint of the chord  $EF$ .  
 2. \*\*Understand the implications of the midpoint:\*\*\n - Since  $G$  is the midpoint of  $EF$ ,  $CG$  is perpendicular to  $EF$ . This means that  $\angle CGE = 90^\circ$ .  
 3. \*\*Use the inscribed angle theorem:\*\*\n - The angle subtended by an arc at the center of the circle is twice the angle subtended by the same arc at any point on the circumference.  
 4. \*\*Determine the measure of  $\angle ECF$ :\*\*\n - Since  $\angle DCF = 20^\circ$  and  $\angle CGE = 90^\circ$ , we can find  $\angle ECF$  as follows:  

$$\angle ECF = 90^\circ - 20^\circ = 70^\circ$$
  
 5. \*\*Calculate  $\angle EOD$ :\*\*\n - Using the relationship between the central angle and the inscribed angle:  

$$\angle EOD = 2 \times \angle ECF = 2 \times 70^\circ = 140^\circ$$
  
 Thus, the measure of angle  $EOD$  is  $140^\circ$ .  
 </think><answer>140</answer>

Figure 16: Comparison between TON and vanilla GRPO in GeoQA

Q: In the provided figure, the city government is making plans to construct an overpass with a height AB measuring 6.0. If angle ACB represents the angle between the slope AC of the overpass and the ground BC, and the sine of angle ACB is equal to 3.0/5.0, what is the length of slope AC?



TON:

<think>The length of the slope AC can be calculated using the Pythagorean theorem. Since the ratio is 3:4, we can set up a right triangle where the height (AB) is 6.0 units and the horizontal distance is 8.0 units. Since the ratio is 3:4, we can use this to find the length of BC.

1. **Identify the ratio and relate it to the sides of the triangle:**

- Let the length of AB be  $h$  and the length of BC be  $d$ .

- Given  $AB = 6.0$  units, and the ratio is 3:4, we can express  $d$  in terms of  $h$ :

$$\frac{h}{d} = \frac{3}{4}$$

Solving for  $d$ :

$$d = \frac{4}{3} \times 6.0 = 8.0 \text{ units}$$

2. **Apply the Pythagorean theorem to find the hypotenuse (AC):**

$$AC = \sqrt{AB^2 + BC^2}$$

Substituting the known values:

$$AC = \sqrt{6.0^2 + 8.0^2}$$

Calculate the squares:

$$6.0^2 = 36.0 \quad \text{and} \quad 8.0^2 = 64.0$$

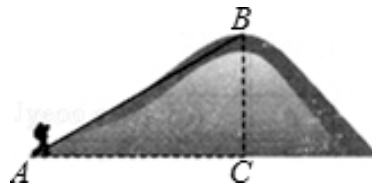
Add these values:

$$AC = \sqrt{36.0 + 64.0} = \sqrt{100.0} = 10.0 \text{ units}$$

Therefore, the length of the slope AC is 10.0 units.</think><answer>10</answer>

Figure 17: Comparison between TON and vanilla GRPO in GeoQA

Q: In the provided figure, the city government is making plans to construct an overpass with a height AB measuring 6.0. If angle ACB represents the angle between the slope AC of the overpass and the ground BC, and the sine of angle ACB is equal to  $\frac{3.0}{5.0}$ , what is the length of slope AC?



GRPO

<think>

To solve for the length of the slope AC, we can use the given information about the slope and the vertical height. The slope AC is in the ratio 3.0:4.0, and the vertical height AB is 6.0 units.

1. **Identify the components of the right triangle:**

- The vertical height (AB) is 6.0 units.
- The ratio of the horizontal length (BC) to the vertical length (AB) is 3.0:4.0, so we can denote BC as  $\frac{3}{4} \times AB$ .

2. **Calculate BC:**

- $BC = \frac{3}{4} \times 6.0 = 4.5$  units.

3. **Use the Pythagorean theorem to find AC:**

- The Pythagorean theorem states  $AC^2 = AB^2 + BC^2$ .
- Substitute the known values:  $AC^2 = 6.0^2 + 4.5^2$ .
- $AC^2 = 36.0 + 20.25 = 56.25$ .
- $AC = \sqrt{56.25} = 7.5$  units.

Therefore, the length of the slope AC is 7.5 units.

</think>

<answer>

7.5

</answer>

Figure 18: Comparison between TON and vanilla GRPO in GeoQA