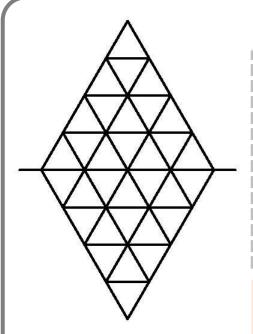
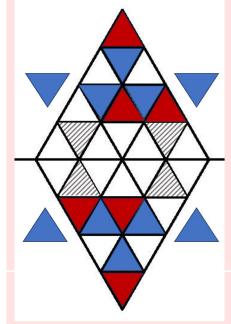
## Math: 2D Transformation



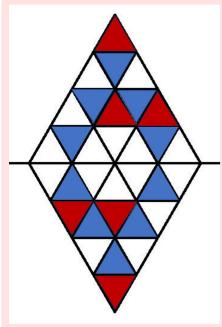


**Question:** Each half of this figure is composed of 3 red triangles, 5 blue triangles and 8 white triangles. When the upper half is folded down over the centerline, 2 pairs of red triangles coincide, as do 3 pairs of blue triangles. There are 2 red-white pairs. How many white pairs coincide?

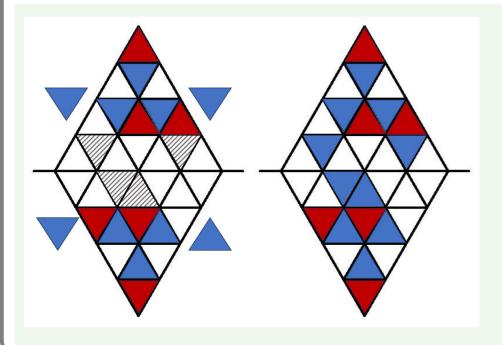
## Error Analysis:



**O1:** 1.Count the total number of triangles per half: ... 2. When folded: ... 3. Given pairings: ... 4. Account for these given pairs: ... 5. Update the counts after using these pairs: ... 6. Remaining pairs to form: ... The remaining top and bottom sets are the same: 2B + 7W on top and 2B + 7W on bottom. 7. Forming the remaining pairs: With equal numbers of blues and whites on both sides, the most natural arrangement is to pair like colors with like colors: Pair the 2 remaining top B with the 2 remaining bottom  $B \rightarrow 2$  more B-B pairs. - Pair the 7 remaining top W with the 7 remaining bottom  $W \rightarrow 7$  W-W pairs. 8. ... Final Answer: \(\\boxed{7}\\)



The model reasoning suffers from two key issues. First, it failed to "remember" prior pairings, which humans would track easily using the diagram, revealing its inability to manage dynamic states in multi-step tasks without visual feedback. Second, tasks without visual feedback. Second, it ignored image-derived constraints, like the "3-blue pairs" rule, and wrongly assumed pairing like colors was natural. This highlights the model's broader limitation in integrating textual and visual information, particularly for tasks requiring spatial reasoning and multi-hop logic, where visual aids simplify the solution.



## **Human Reasoning (Ground Truth):**

Forming the remaining 2 top B and 2 bottom B, the image shows there are already 3 B-B pairs. Thus, the 2 top B and 2 bottom B cannot form more B-B pairs but instead create 4 B-W pairs. After this, the remaining W-W pairs total 5, as counted in the image.