Challenge 1: Reasoning-Action Inconsistency

Driver

Turn Right

AR-VLM

The most appropriate meta behavior is to **speed up** and **turn right**.

Sequential Generation

Suggested Trajectory: [[0.0,0.0],[0.0,0.0],[0.0,0.0],[0.01,0.0],[0.53,0.0]]

Driver

Turn Right

dVLM

The most appropriate meta behavior is to ____ and ____.
Suggested Trajectory: [[_,_],[_,_],[0.0,0.0],[_,_],[0.0,_]]

Iteratively Refining

The most appropriate meta behavior is to stop and keep lane.
Suggested Trajectory: [[0.0,0.0],[0.0,0.0],[0.0,0.0],[0.0,0.0],[0.0,0.0]]

Challenge 2: Uncontrollable Generation

Driver

Go Straight. Output trajectory only

AR-VLM

The ego vehicle is expected to decelerate and prepare to stop due to the red light…

Bypass!

Suggested Trajectory: [[8.2,0.1], [15.5,0.2], [21.8,0.4], [27.0,0.5], [31.1,0.5]]

Driver

Go Straight. Output trajectory only

Fixed CoT Template

dVLM

Scene Reasoning: _____.
Meta Behavior: ____ and ____.
Suggested Trajectory: _____.

Scene Reasoning: The ego vehicle is expected to decelerate and prepare to stop due to the red light…
Meta Behavior: stop and straight.
Suggested Trajectory: [[1.3,0],[1.3,0],[1.3,0],[1.3,0],[1.3,0]].

Fill in Blank