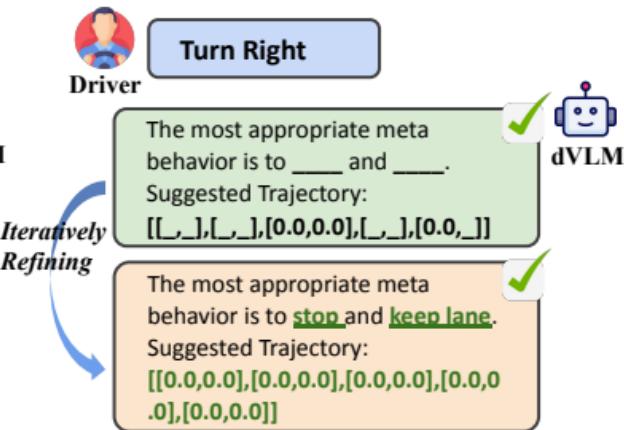
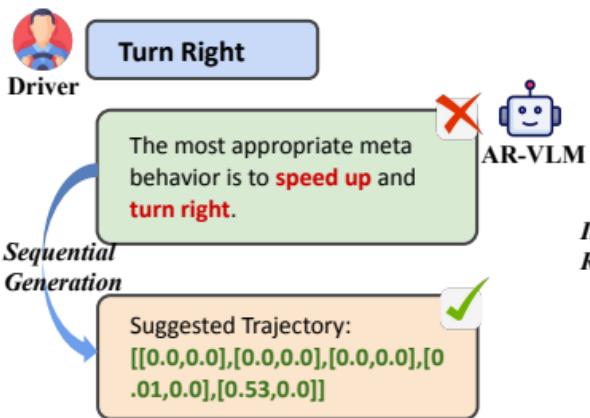


Challenge 1: Reasoning-Action Inconsistency



Challenge 2: Uncontrollable Generation

