Automated Refinement with Uncertainty Disentanglement \square (car $\land \neg green_light \rightarrow \neg turn_right$) **Perception Uncertainty** \square (pedestrian \rightarrow wait) $u_p \ge t_p$ u_p Temporal Logic Specifications Image Set Task Bank Model Checking nerato Vision Wait for **Query with** pedestrian, then **Text Prompt** turn right Plan Embedding Input Image Automaton **Foundation Model Fine-Tuning Data** { input: [image, text prompt], **Fine-Tune and Re-Calibrate** Output: [plan]}