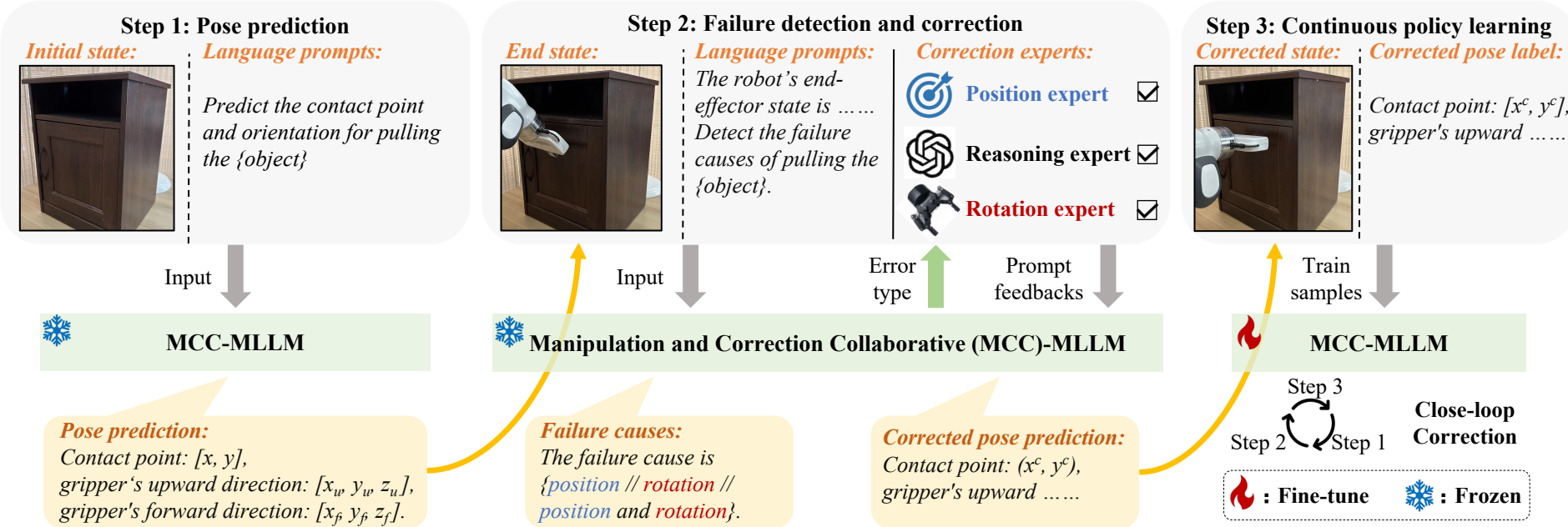
Step 1: Pose prediction
Initial state:
Language prompts:
Predict the contact point and orientation for pulling the {object}
Step 2: Failure detection and correction
End state:
Language prompts:
The robot's end-effector state is …… Detect the failure causes of pulling the {object}.
Correction experts:
Position expert
Reasoning expert
Rotation expert
Step 3: Continuous policy learning
Corrected state:
Corrected pose label:
Contact point: [x^c, y^c], gripper's upward ……
Input
Input
Error type
Prompt feedbacks
Train samples
MCC-MLLM
Manipulation and Correction Collaborative (MCC)-MLLM
MCC-MLLM
Pose prediction:
Contact point: [x, y],
gripper's upward direction: [x_u, y_u, z_u],
gripper's forward direction: [x_f, y_f, z_f].
Failure causes:
The failure cause is {position // rotation // position and rotation}.
Corrected pose prediction:
Contact point: (x^c, y^c),
gripper's upward ……
Step 3
Step 2
Step 1
Close-loop Correction
: Fine-tune
: Frozen