

Setup



1,334 trajectories
VLA's failures, not synthetic



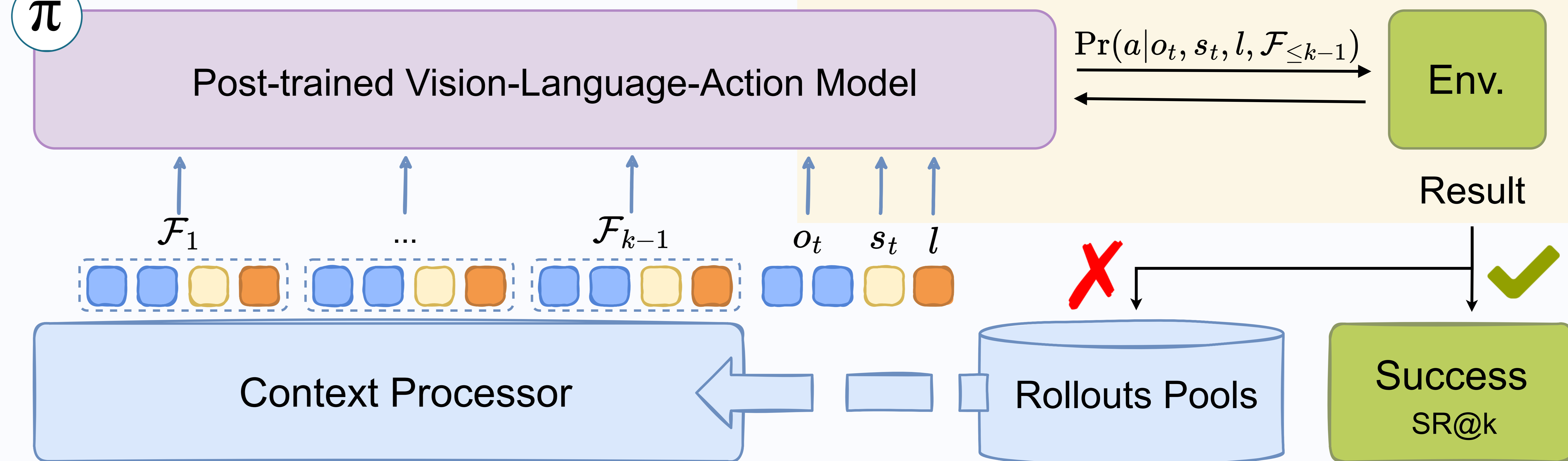
VLA in-context learning
no parameter update



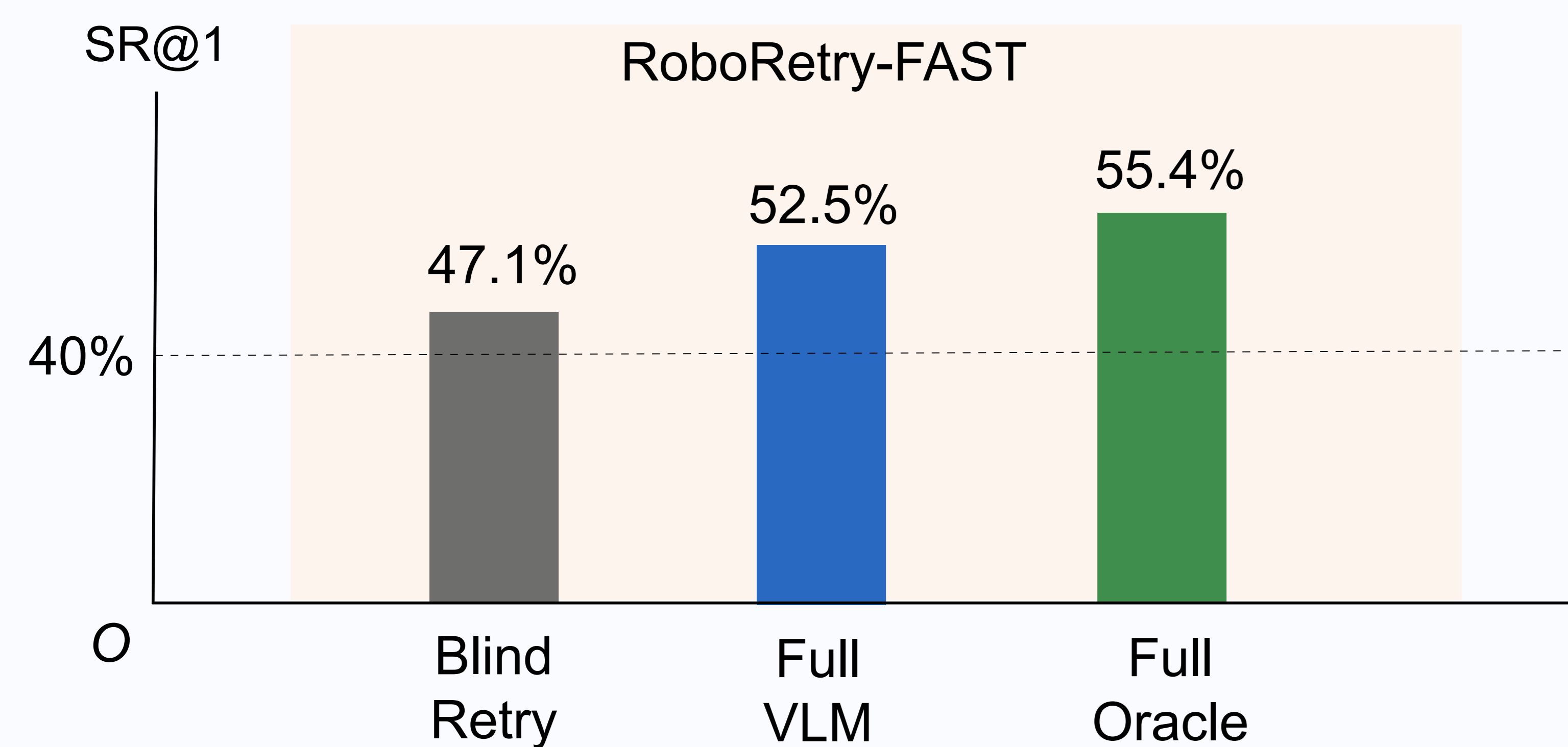
diagnostic analysis
no extra recovery module

Probe

π

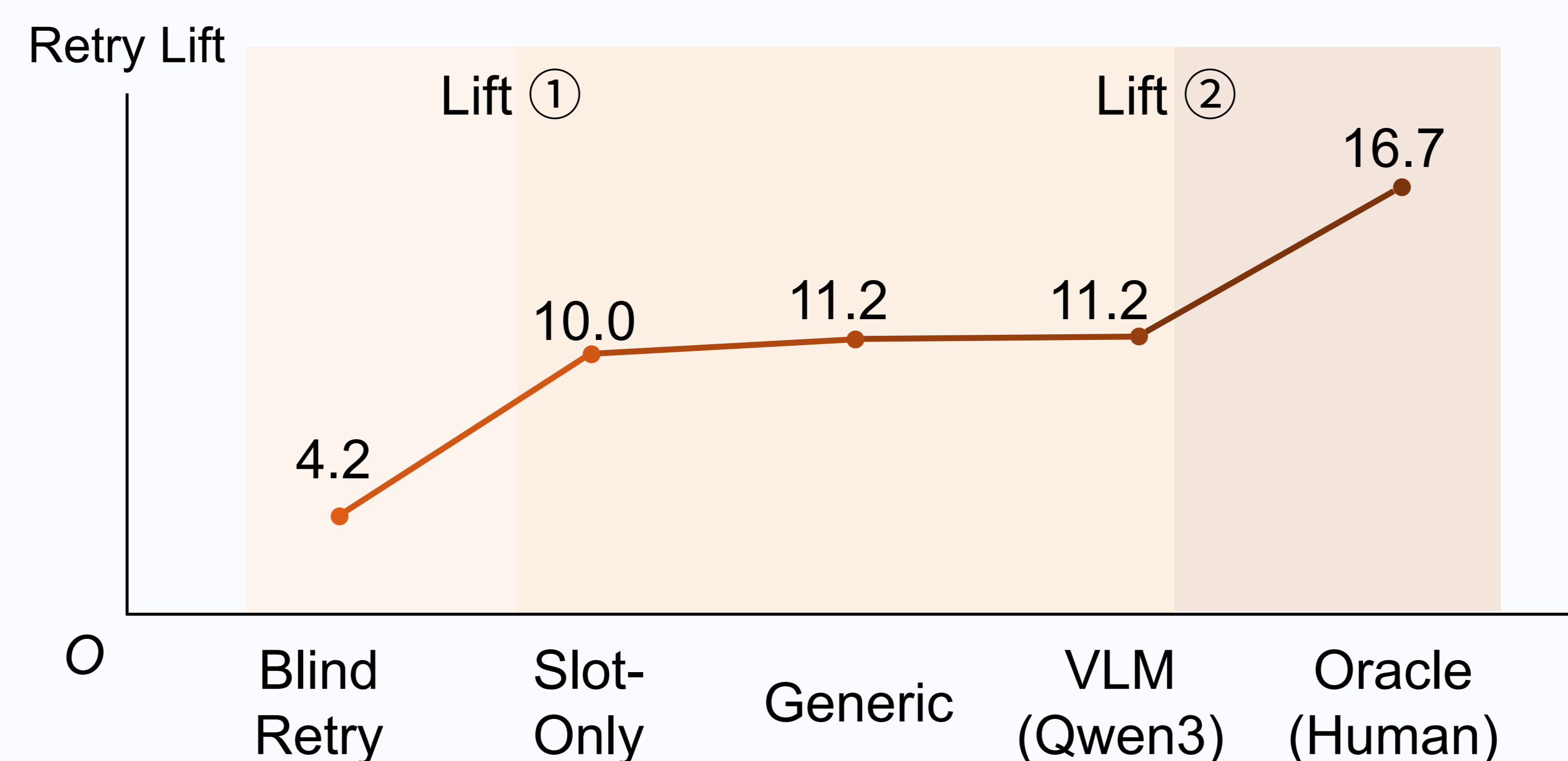


Does failure context improve retry?



Failure slots improve retry beyond blind retry.

How does it improve retry performance?



Retry lift combines a slot-format cue and content-dependent correction.

When is retry still insufficient?

Works well

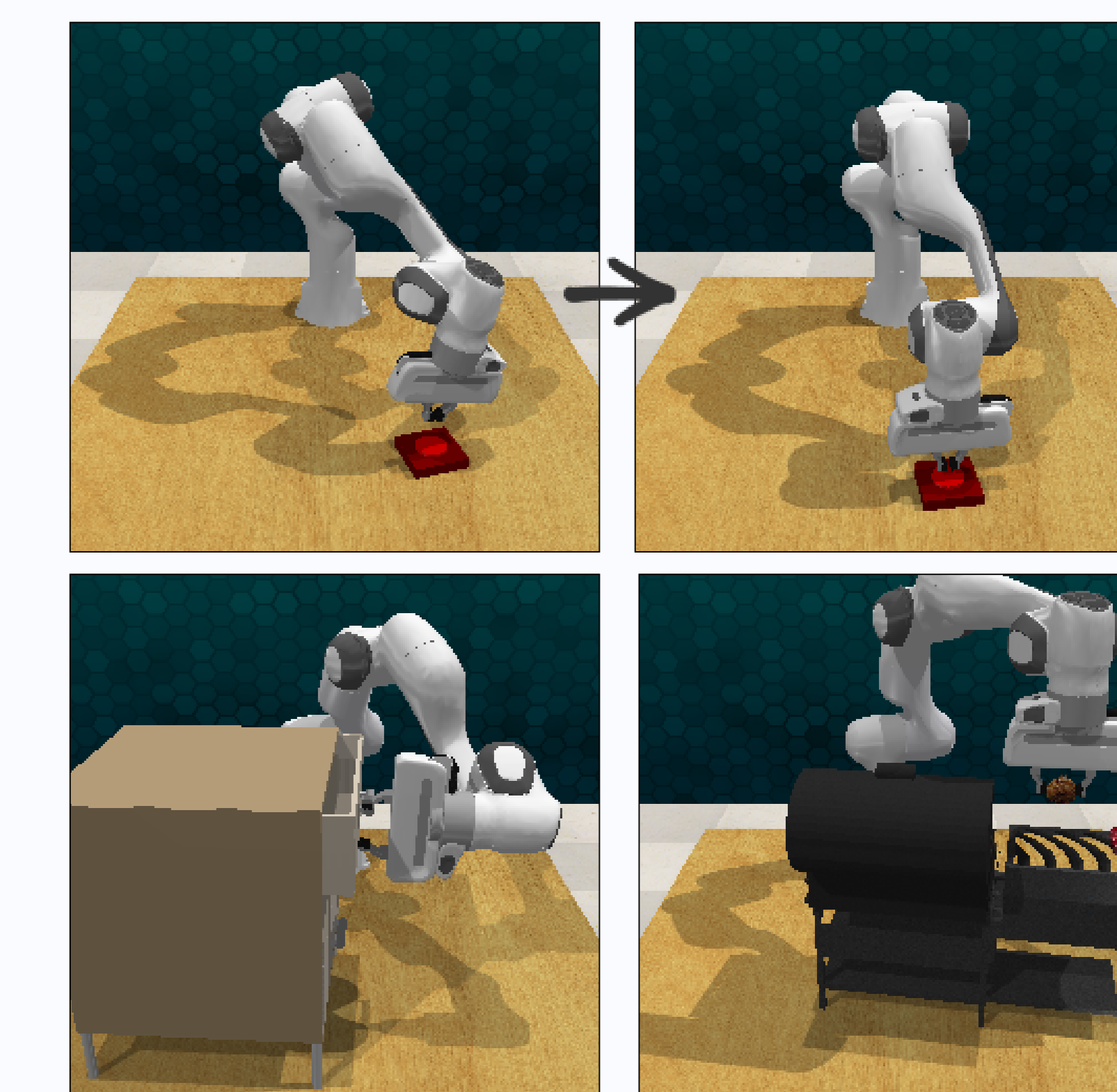
spatial corrections

Still hard

target-regrounding

progress-regrounding

precision-sensitive



A more native and actionable representation is needed for future