SafeDec: Constrained Decoding for Safe Autoregressive Generalist Robot Navigation Policies

Parv Kapoor, Akila Ganlath, Michael Clifford, Changliu Liu, Sebastian Scherer, Eunsuk Kang

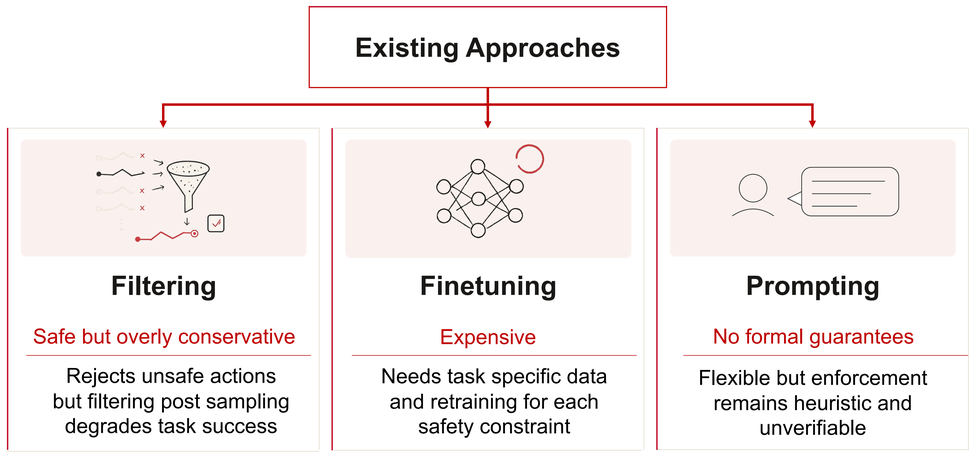
Carnegie Mellon University

ICML International Conference On Machine Learning

1. Generalist robot policies need runtime safety assurance

Generalist robot policies trained on large datasets show impressive generalization but also need to satisfy **behavioral rules for safe deployment**

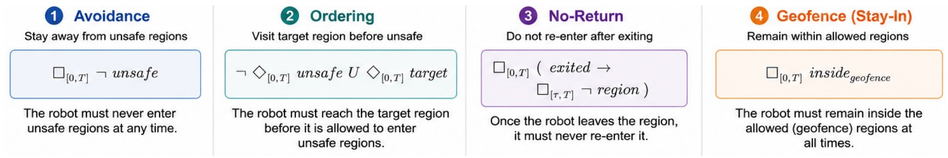
Existing enforcement methods are restrictive, expensive or lack formal guarantees

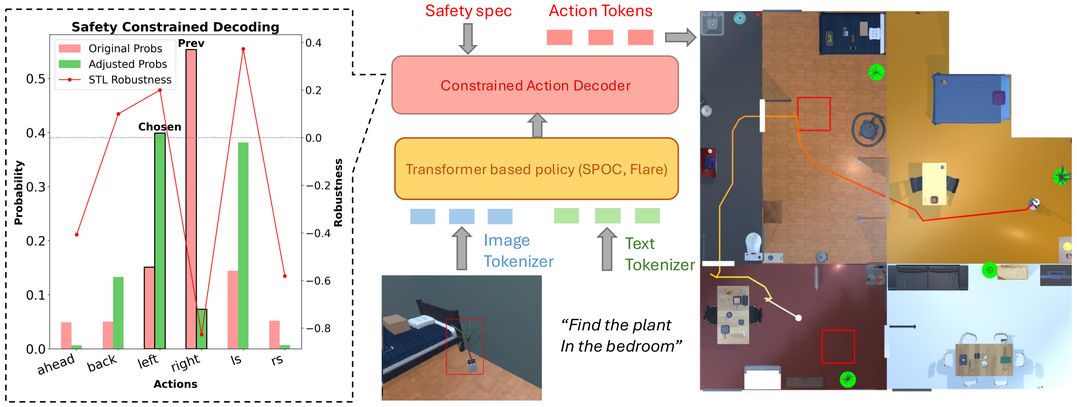


Need: A training-free, adaptable, and formally grounded safety mechanism

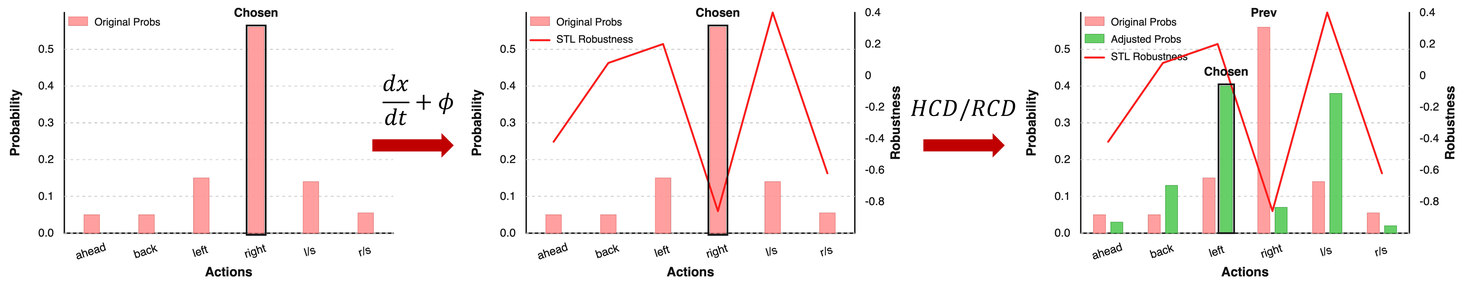
2. SafeDec: Enforcing safety rules while decoding actions

Given multimodal observations, a pretrained transformer-based navigation policy generates candidate actions. Our constrained decoder then filters these actions using satisfaction scores from a user-defined **Signal Temporal Logic (STL)** specification





3. Two decoding strategies: Hard and Robust Constrained Decoding



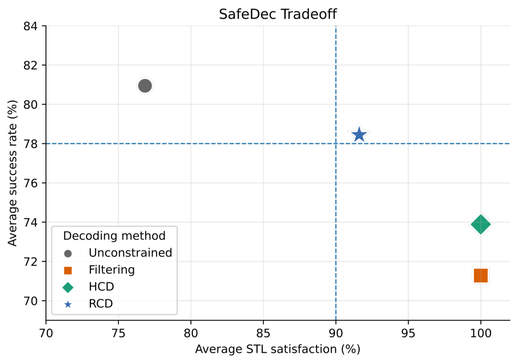
1. Get logits predicted by the model
2. Step through a simple dynamics model
3. Check spec satisfaction based on predicted states

HCD
If violated, set logits to $-\infty$ and resample

RCD
If violated, reweight actions by ρ and resample

4. SafeDec ensures specification satisfaction across various policies!

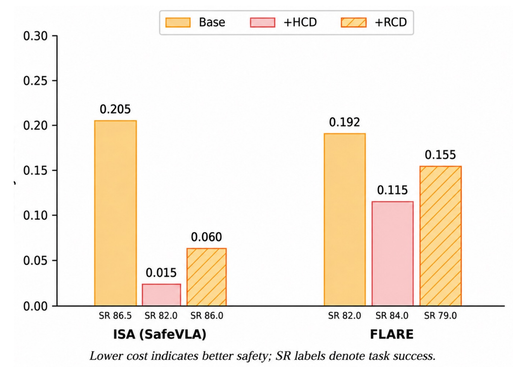
Tested across 3 state-of-the-art pretrained policies (SPOC, PoliFormer and Flare) across 200 procedurally generated environments in AI2-THOR for object & room navigation tasks



Baselines
Unconstrained: Base robot policy
Filtering: Post-decoding enforcement

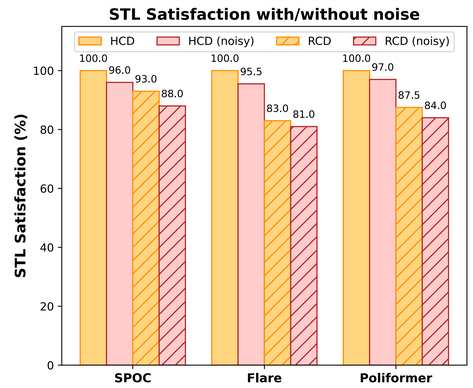
HCD vs Filtering
Both reach 100% STL satisfaction
HCD recovers +0.5–3 pts of task success

RCD: Trade-off between safety & task succ.
RCD drops STL satisfaction but stays within 0.5–6.5 pts of original task success



Lower cost indicates better safety; SR labels denote task success.

SafeDec substantially reduces safety cost for both Safe RL policies



SafeDec maintains high satisfaction even under noisy dynamics making it a robust safety enforcement mechanism

SafeDec **improves safety at inference time** without retraining or modifying the base robot policy

HCD gives strict STL satisfaction; RCD allows **trading off task success vs safety**

Works across SPOC, PoliFormer, Flare, and SafeVLA-style policies, with graceful degradation

Check out **SafeDec**