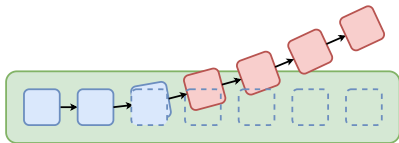




### Curse of Horizon (Guided Decoding)



Token-level control leads to overly long horizons, making alignment fragile.



Generated span



Deviating span



Promising state



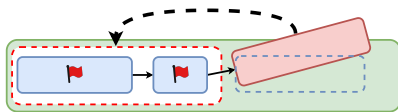
Multi-step  
reward evaluation



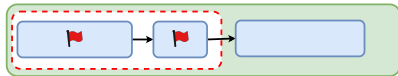
Iterative rollouts



Textual MPC



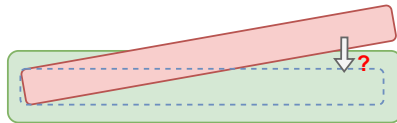
**Hindsight Subgoal Identification**



**Subgoal-Conditioned Re-Generation**



### Curse of Dimensionality (Iterative Refinement)



Response-level rewriting operates in a large action space, causing unstable improvements.

