

Model Input

Prompt

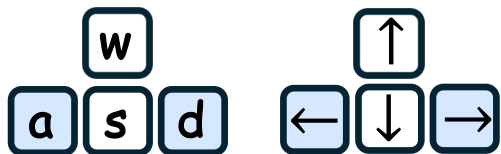
In the **bedroom**,
a **cat** jumps onto the bed.

Expected World State

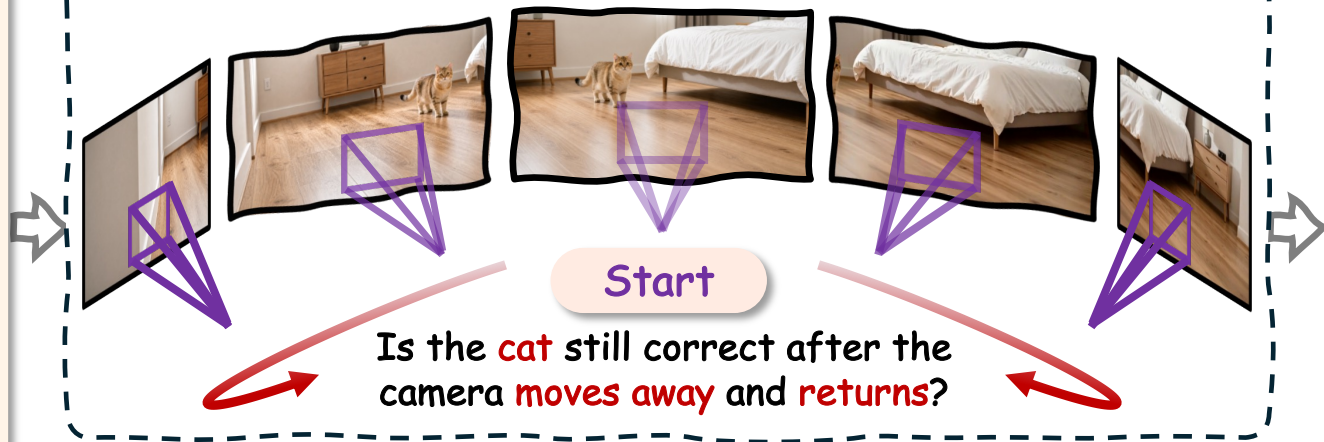
First Frame



Camera Action



Interactive World

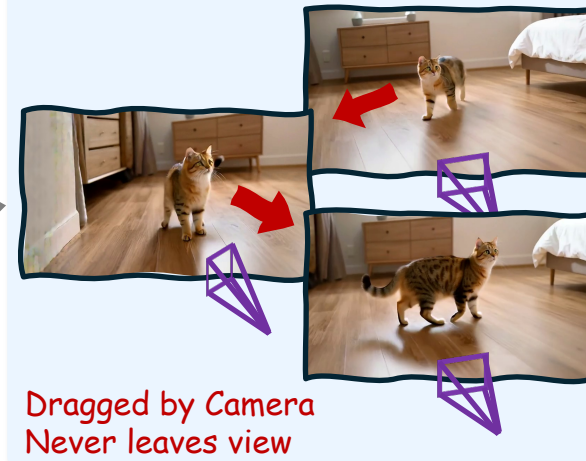


Expected Video



No Robust Evolution

(when camera moves away)



No Returned Consistency

(when camera returns)

