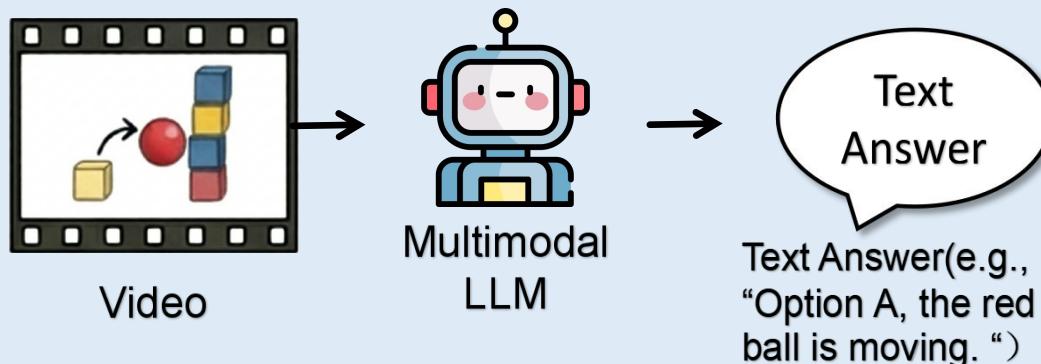
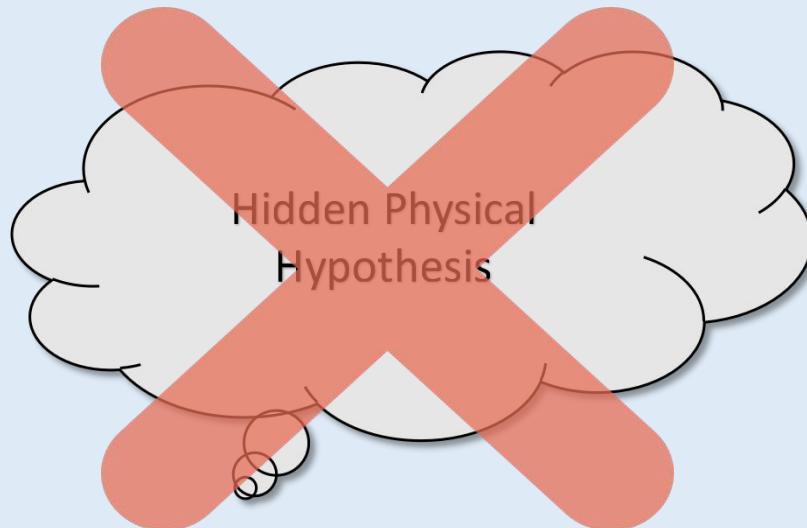


Traditional VQA Paradigm



Evaluates implicit, unverifiable reasoning.



VisPhyWorld Paradigm (Executable Hypothesis)

