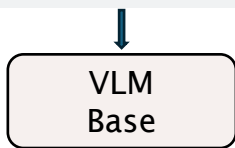
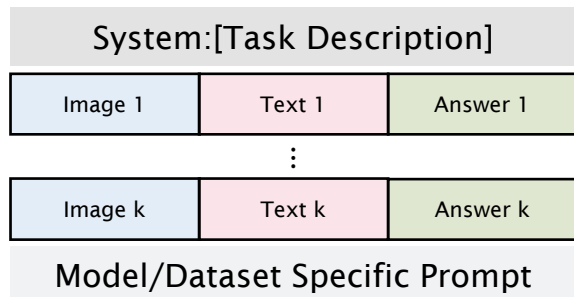
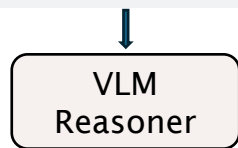
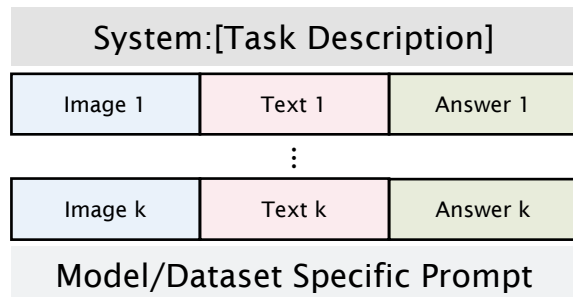


Correctness filter

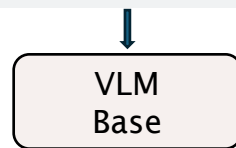
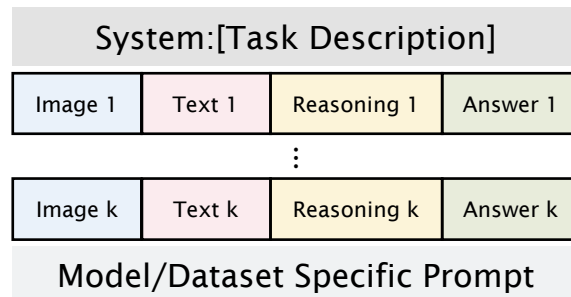
Base Model w/o Reasoning Demos



Reasoning Model w/o Reasoning Demos



Base Model w/ Reasoning Demos



Reasoning Model w/ Reasoning Demos

