

Multi-task Questions

HotpotQA

StrategyQA

MMLU

Multi-method Prompts

CoT

ReAct

Reflexion

X

Tools

LM
(e.g., GPT-4)

generate

Correct trajectories

CoT trajs

ReAct trajs

Reflexion trajs

Incorrect
(removed)

reformat

(a) Fine-tuning

ReAct-format trajectories

LM

fine-tune

FLM

Test Question
(3+4+5=?)

FLM

Tools

CoT

ReAct

Reflexion

Implicit method selection

generate

Question: 3+4+5=?
Thought: 3+4+5=7+5=12. Answer is 12.
Action: Finish[12]
Observation: reward = 1

CoT trajectory in ReAct format