**Research Question:** *How can we train a general-purpose Reasoning LLM across diverse domains via RL?*
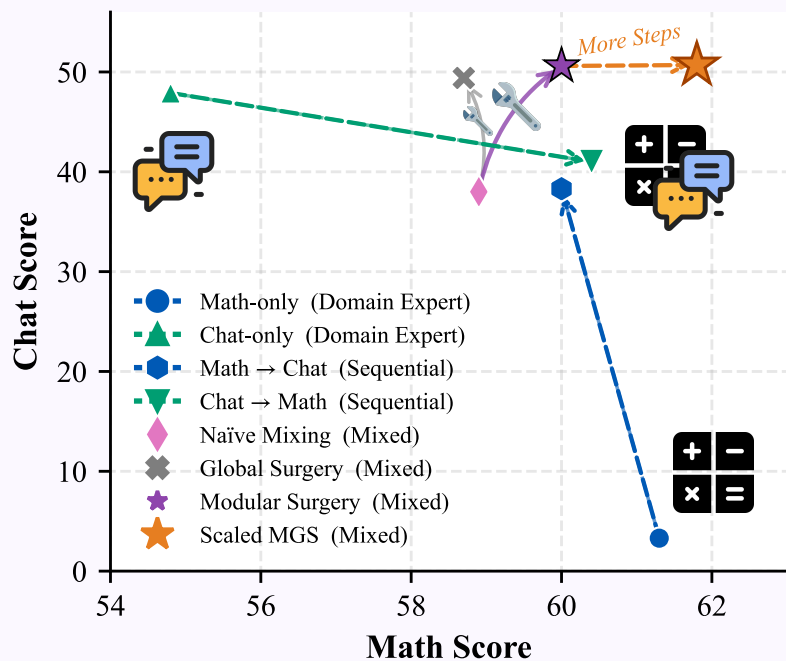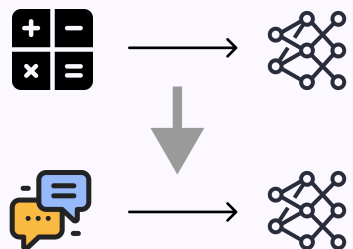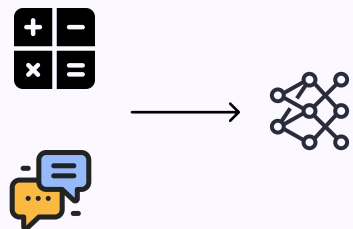


Results Comparison

Sequential RL (Sec. 3.1)

Mixed RL (Sec. 3.3)

Modular Gradient Surgery (Ours, Sec. 4&5)

Conflicting Gradients in each Module

Resolving with Gradient Surgery

Conflicting Task Gradients

Project

Aggregate

More Steps

Math-only (Domain Expert)
Chat-only (Domain Expert)
Math → Chat (Sequential)
Chat → Math (Sequential)
Naïve Mixing (Mixed)
Global Surgery (Mixed)
Modular Surgery (Mixed)
Scaled MGS (Mixed)

Math Score
Chat Score

MLP
Attn
Attn