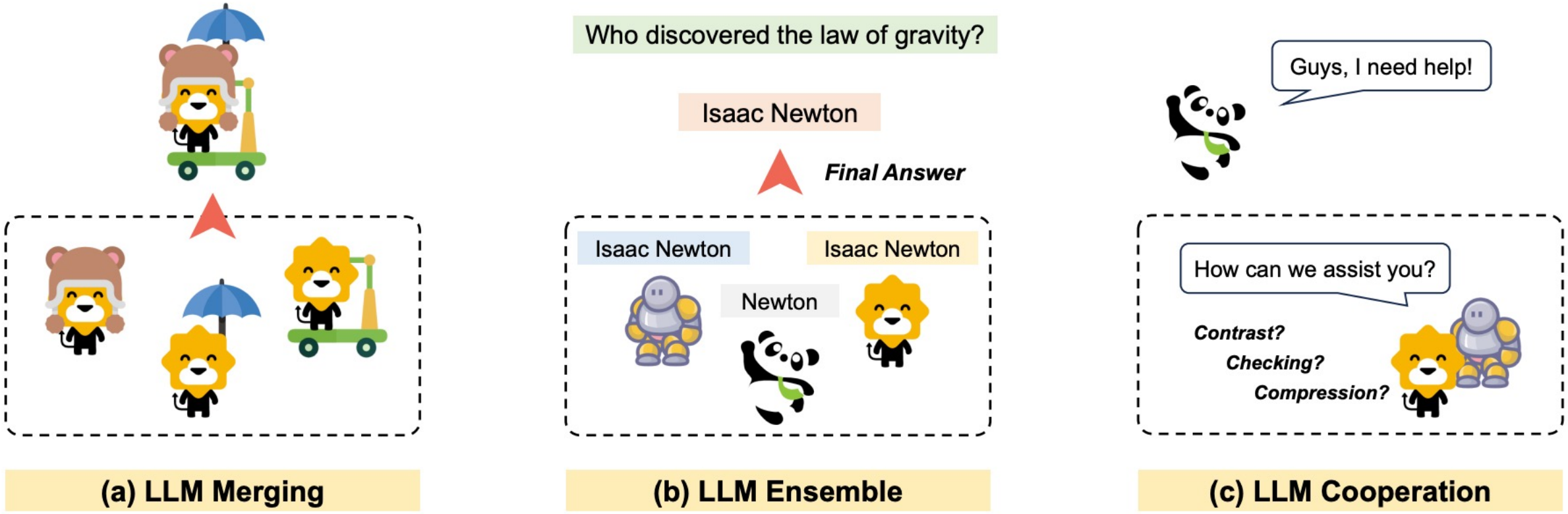
Knowledge Grafting of LLMs

Guodong DU, 25113986R

26 Feb, 2026

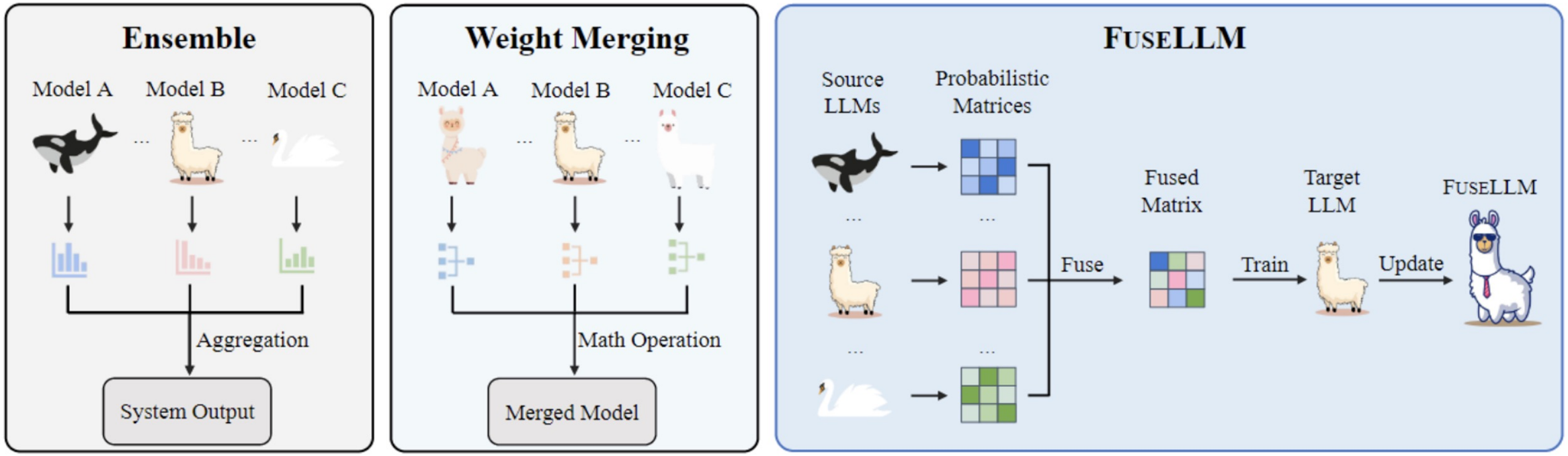
➢ Background: Knowledge Fusion of Large Language Models



Large language models excel at different tasks such as reasoning, dialogue, and domain-specific applications. Can we **combine these strengths** into a smaller, more efficient model?

This challenge is **known as cross-capability transfer**.

➢ What Existing Work Has Done



Two main directions in existing research:

1.Directly combining model parameters to integrate their capabilities.

2.Transferring knowledge from a larger “teacher” model to a smaller “student” model through training.

➢ Research Gap and Motivation

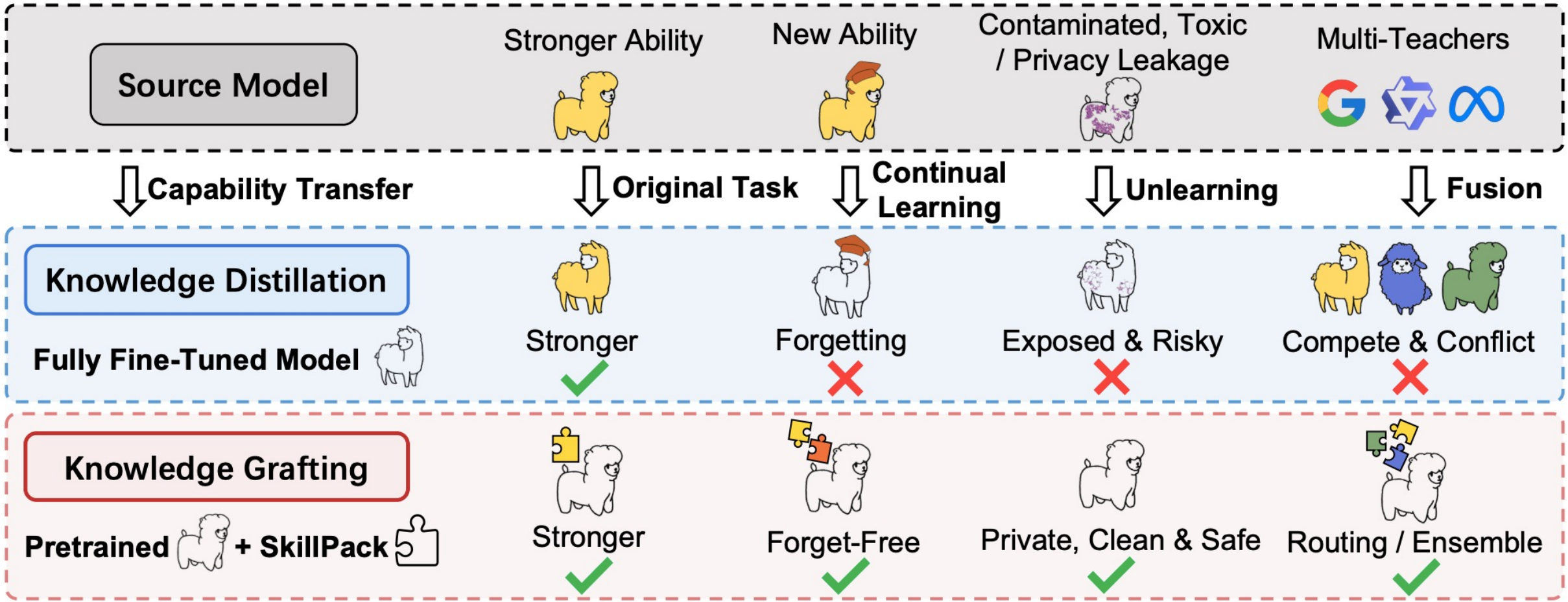


Figure 3: Comparision of knowledge distillation and knowledge grafting in various scenarios.

➢ Research Aim & Proposed Solution

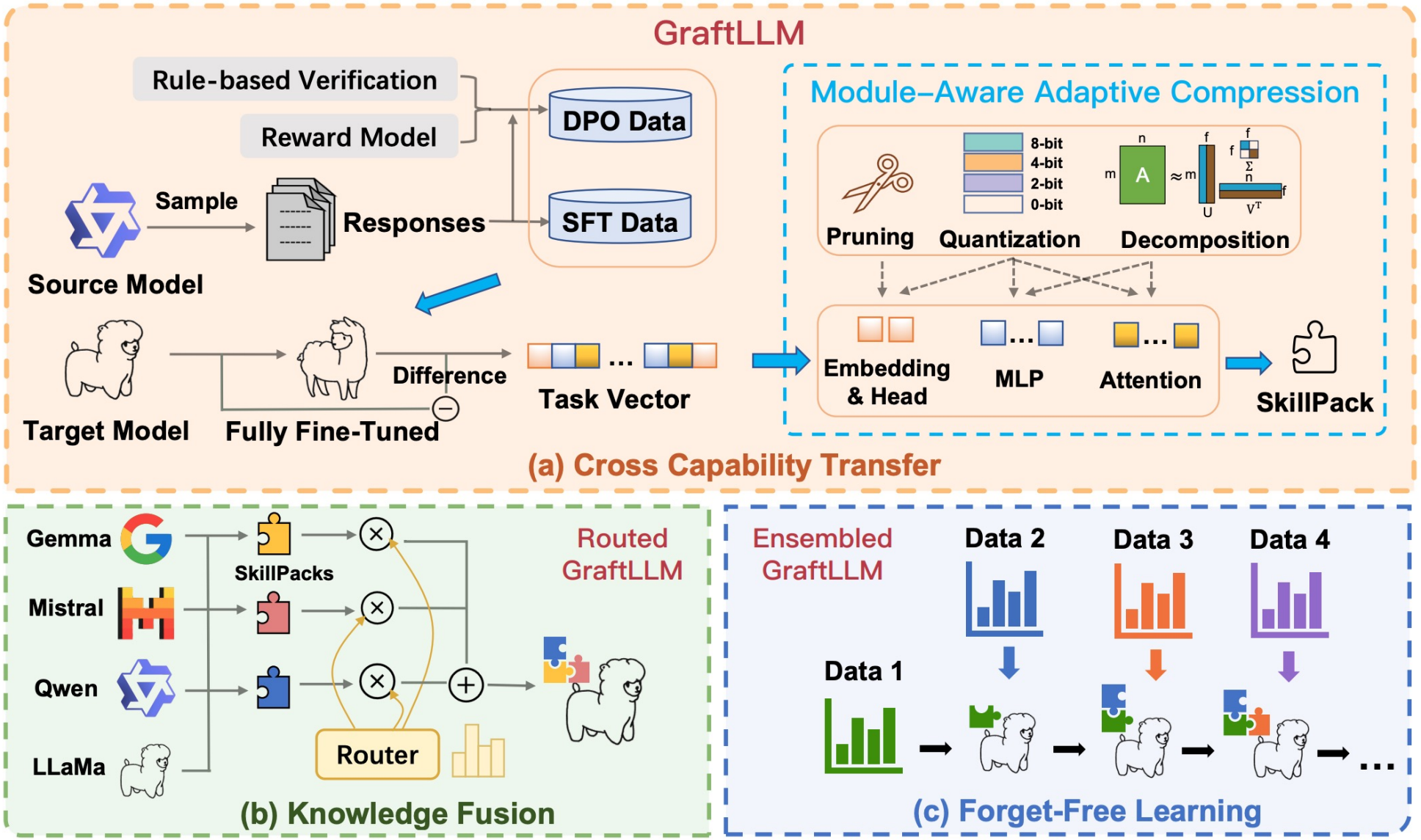


Figure 4: **Overview of GraftLLM**. `GraftLLM` transfers capabilities across heterogeneous LLMs and extracts them into compact SkillPacks, facilitating efficient knowledge fusion and continual learning.

➢ Contribution & Significance

As a result, we enable:

1. Knowledge transfer across heterogeneous models;
2. Knowledge fusion from multiple sources;
3. Forget-free continual learning;

All while keeping parameter storage efficient.