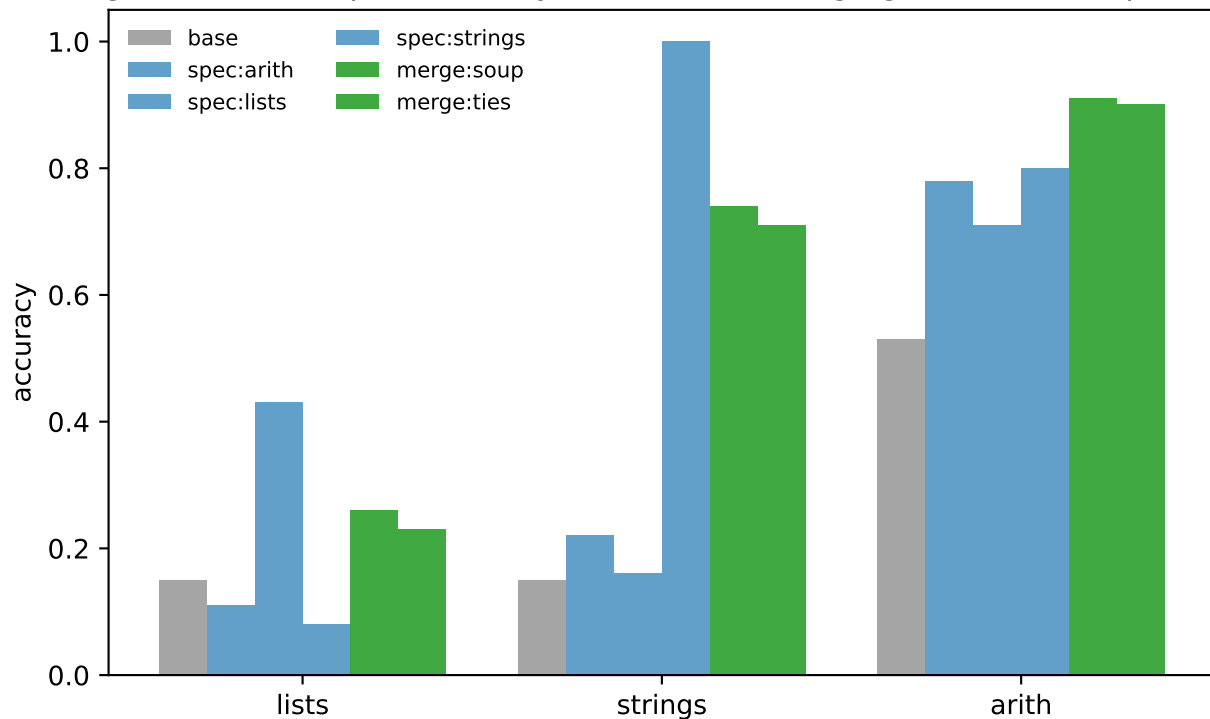


llm_merge — recombining decorrelated specialist LLMs: recombined 0.64 > best specialist 0.63 overall (Qwen2.5-0.5B-Instruct)

Per-family: each specialist spikes on its own family; the merges (green) are competent everywhere (but averaging dilutes some peaks)



Overall (solid) vs worst-family (hatched): the recombined model vs the best single specialist

