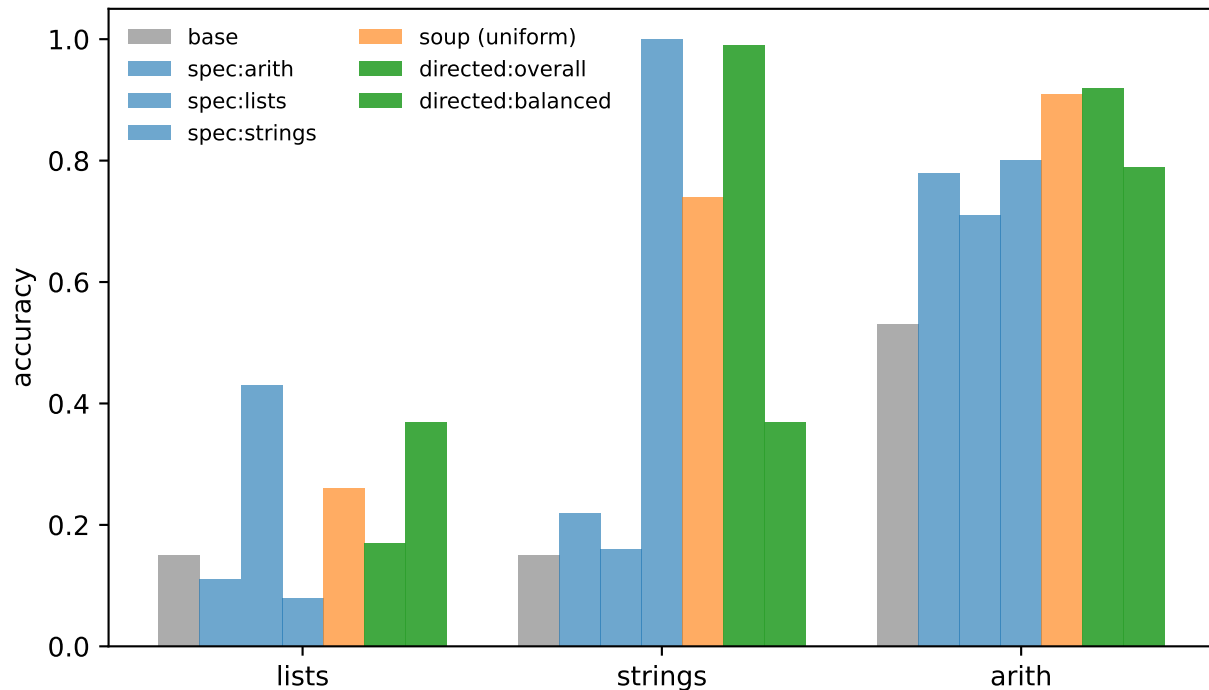


llm_directed — directed sex (breed offspring + select on the verifier): directed 0.69 > soup 0.64 overall (best parent 0.63) (Qwen2.5-0.5B-Instruct)

Per-family: directed offspring (green), selected on the verifier, vs the single uniform soup (orange) and the parents



Overall (solid) vs worst-family (hatched): directed offspring vs soup vs the best parent (dotted)

