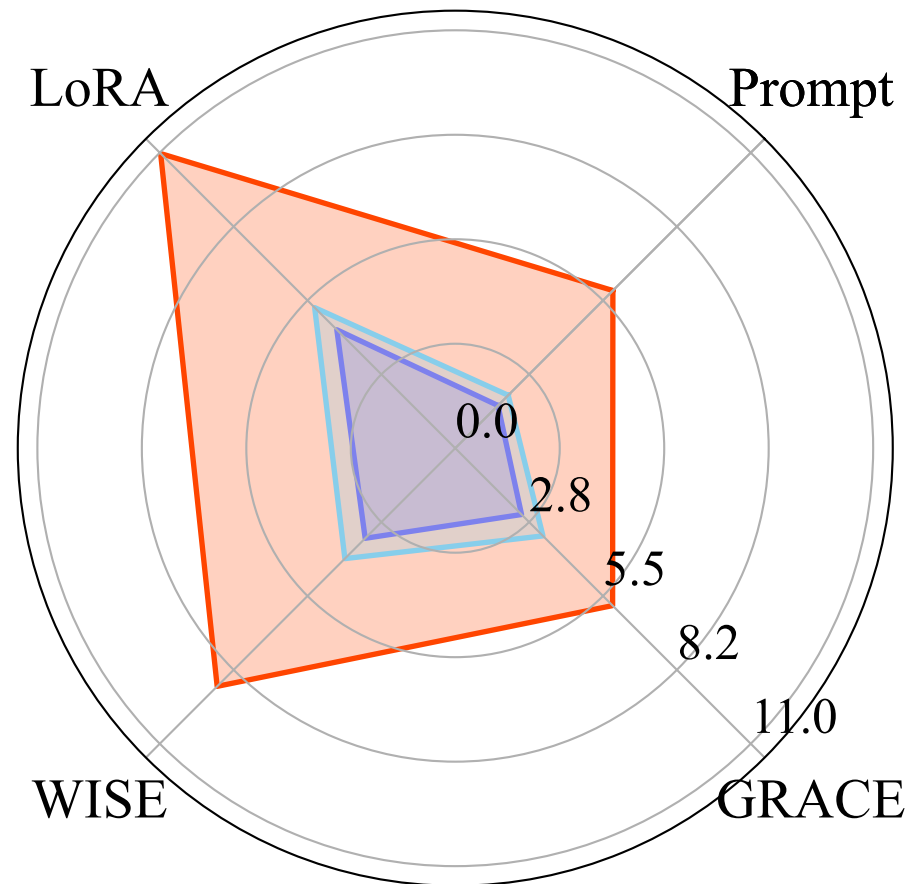
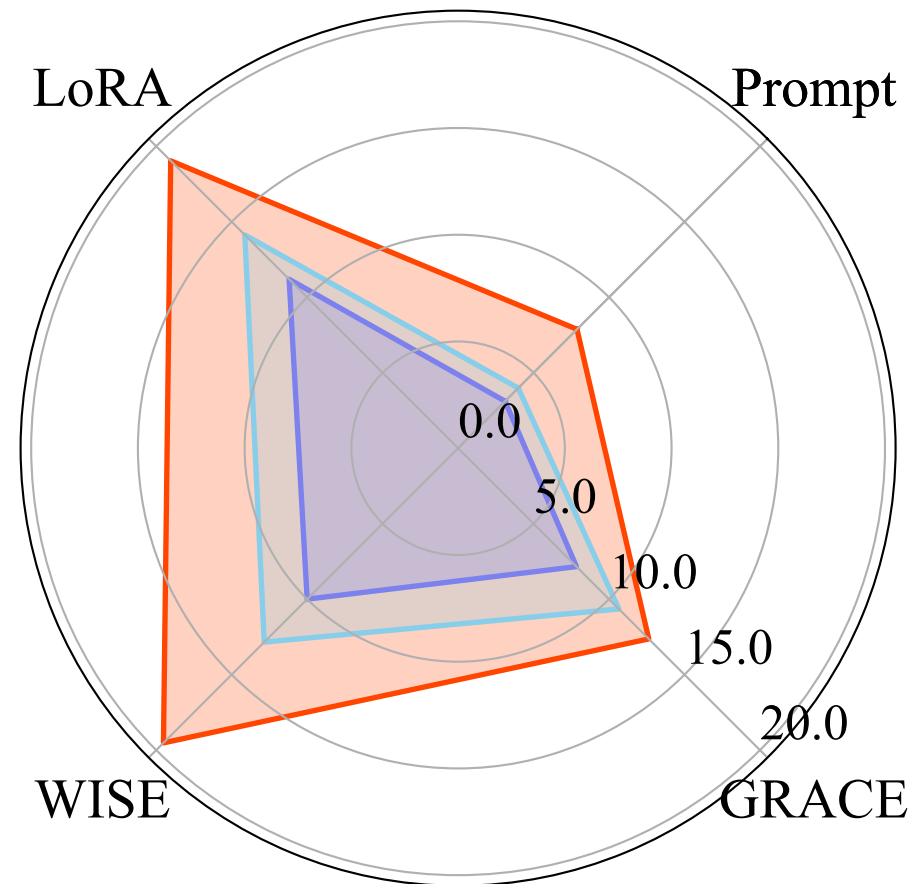


**Avg Time per Sample (s) - Understanding**



**Avg Time per Sample (s) - Reasoning**



— LLaVA-Onevision

— QWen2-VL

— HuaTuoGPT-7B