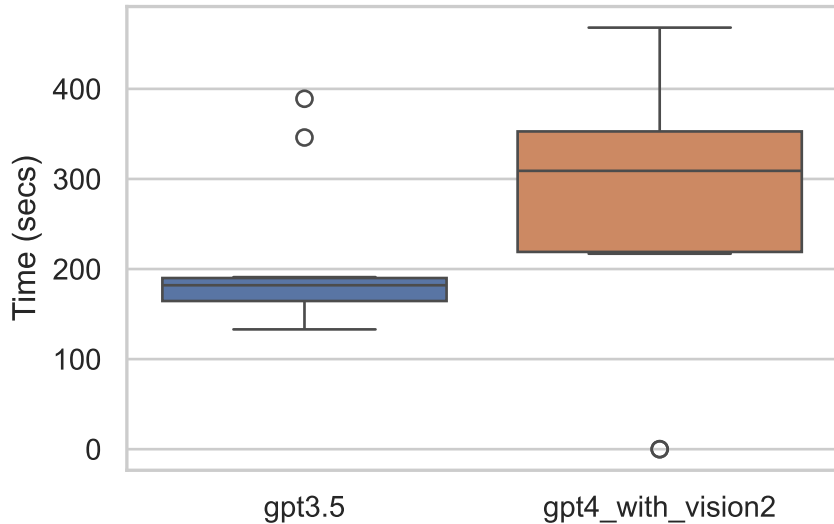
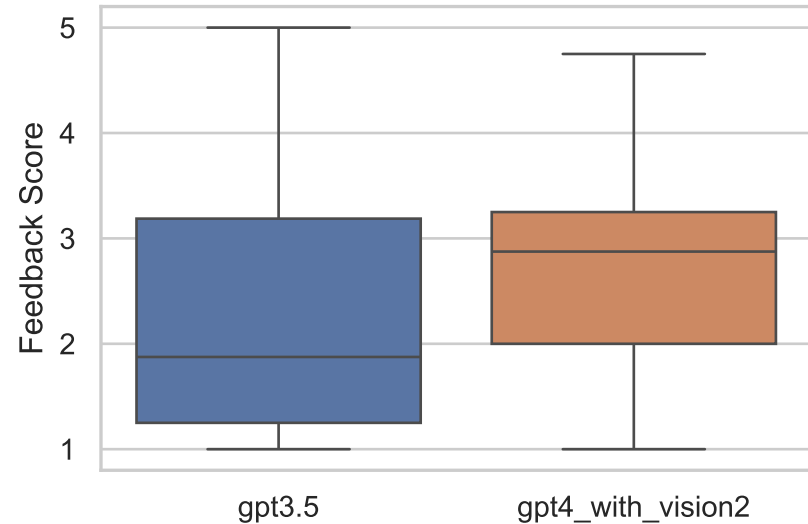


Evaluation of the Multiagent Development Framework

Average Execution Time per Model



Human Feedback per Model



Average Test Iterations for Different Layers

