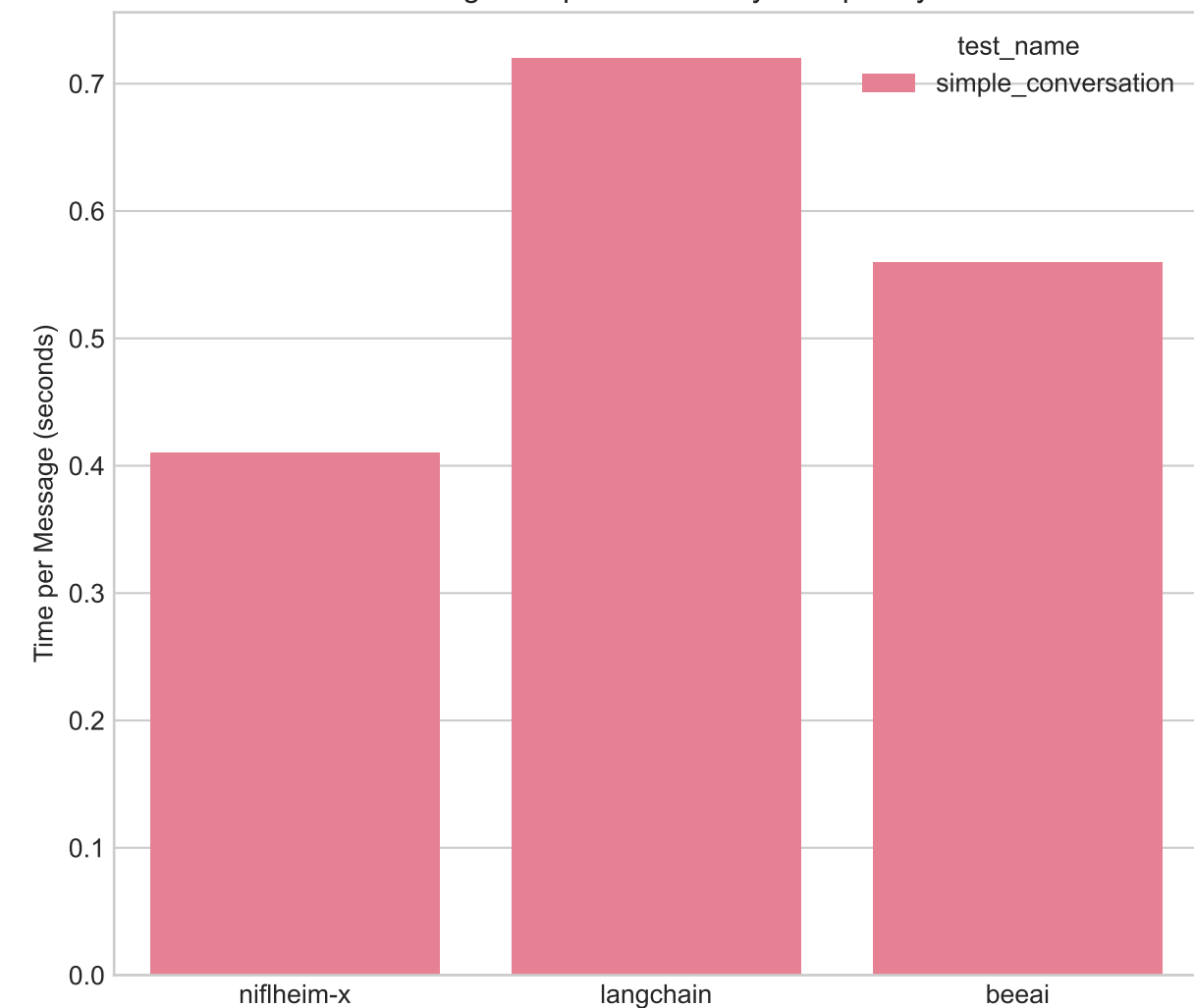
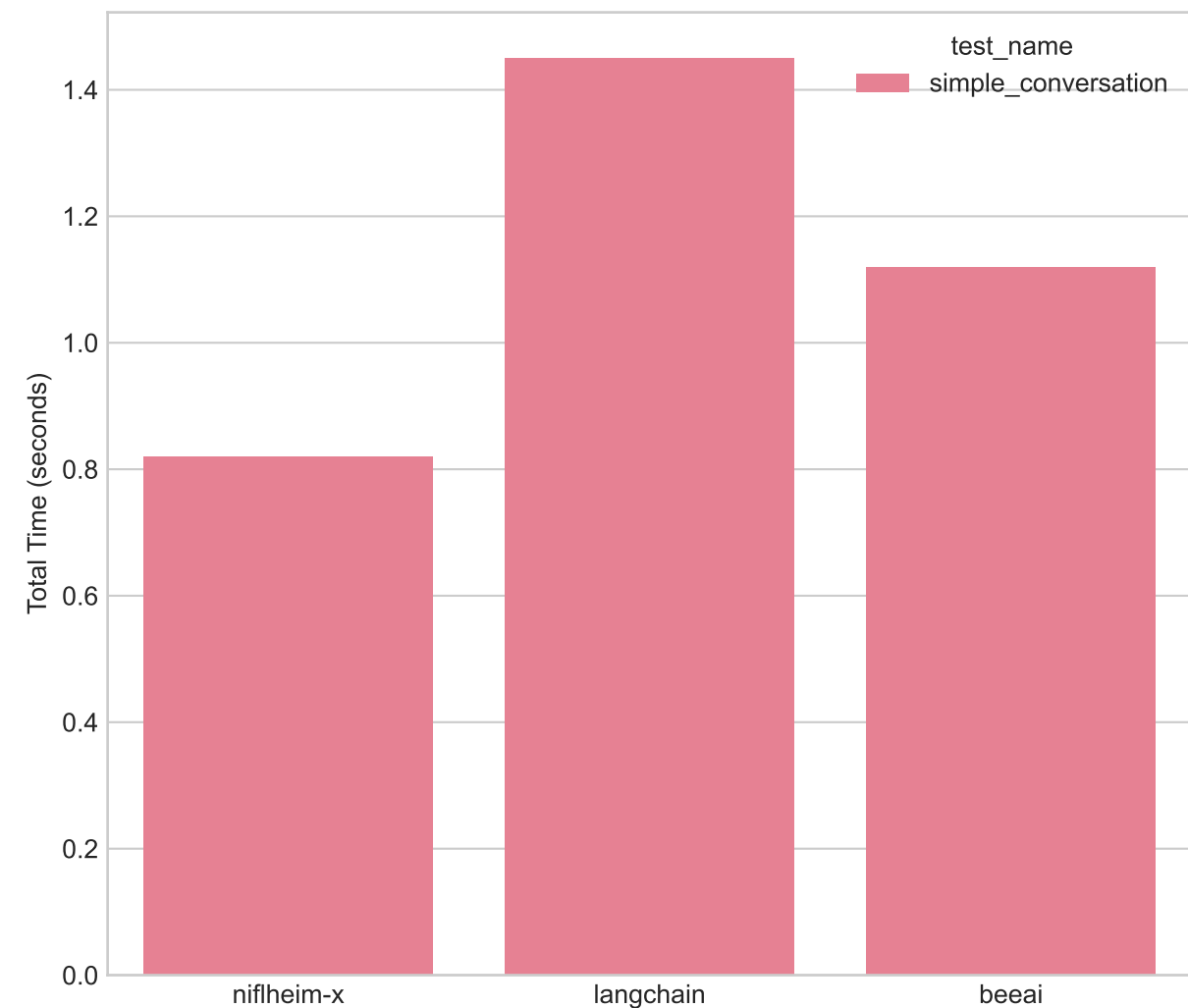


Conversation Performance Comparison

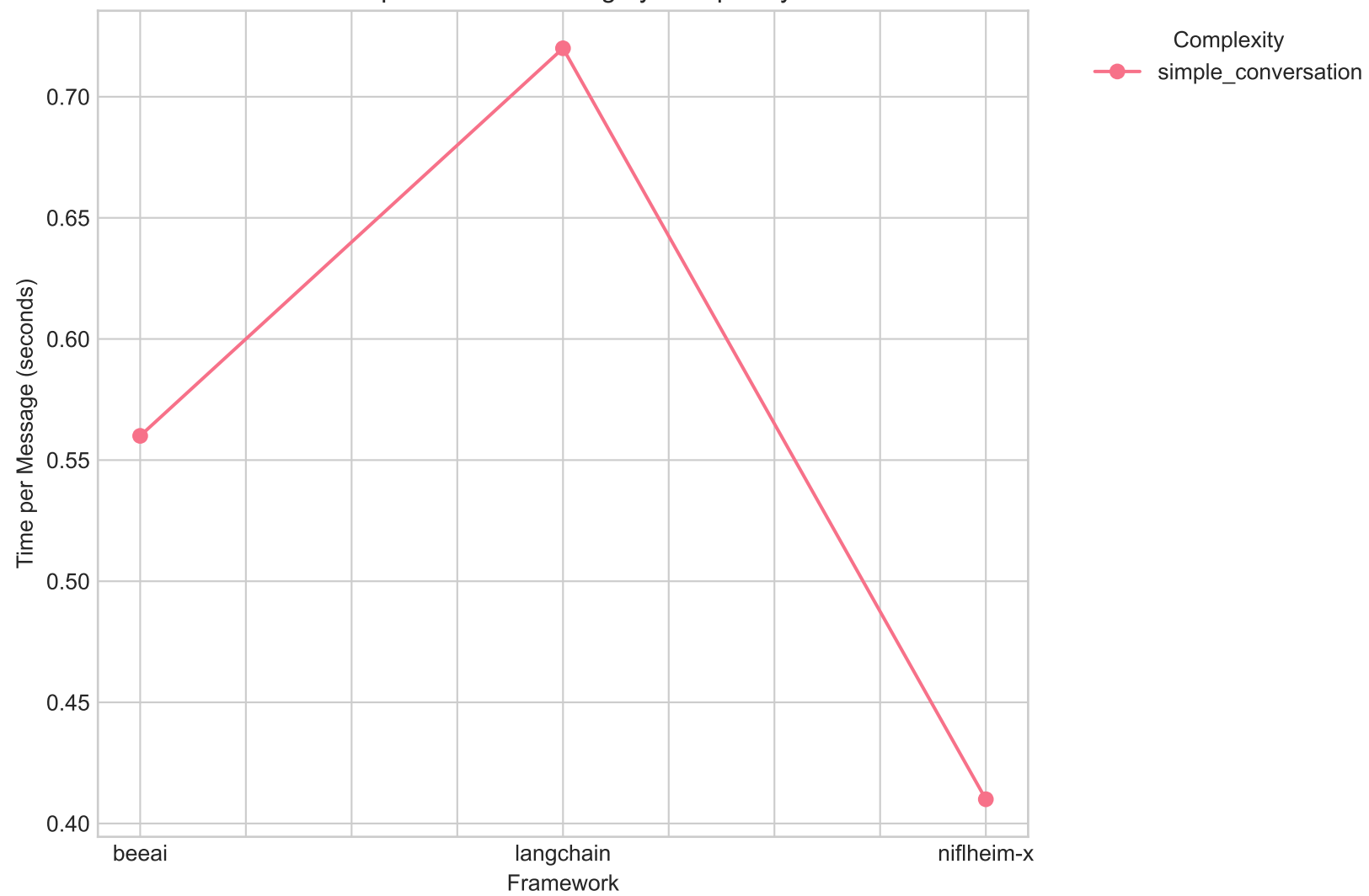
Average Response Time by Complexity



Total Conversation Time



Response Time Scaling by Complexity



Overall Efficiency Score (Lower = Better)

