ShortestPath Task Performance by Path Length and Model (Solid: Avg over Formats, Dashed: Best Format) 0.8 Model 0.7 claude-3.5-sonnet-v2 Average Accuracy gemini-1.5-flash gpt-4o-mini llama3.2-1b-instruct llama3.3-70b-instruct nova-lite nova-pro ---- claude-3.5-sonnet-v2 (Best: ISON-LD) - gemini-1.5-flash (Best: Structured JSON) 0.2 -■- gpt-4o-mini (Best: List of Edges) 0.1 llama3.2-1b-instruct (Best: JSON-LD)

