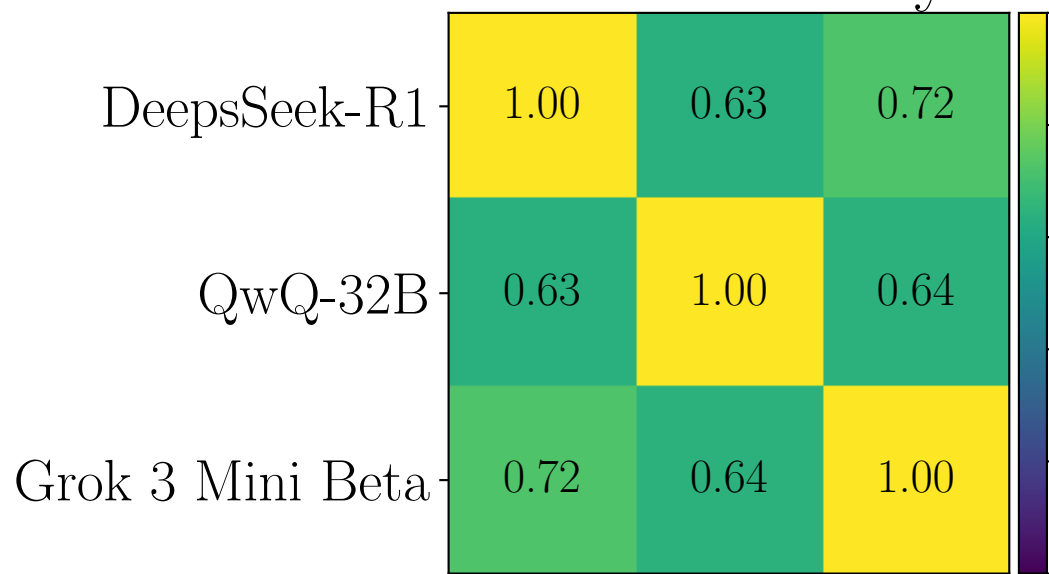
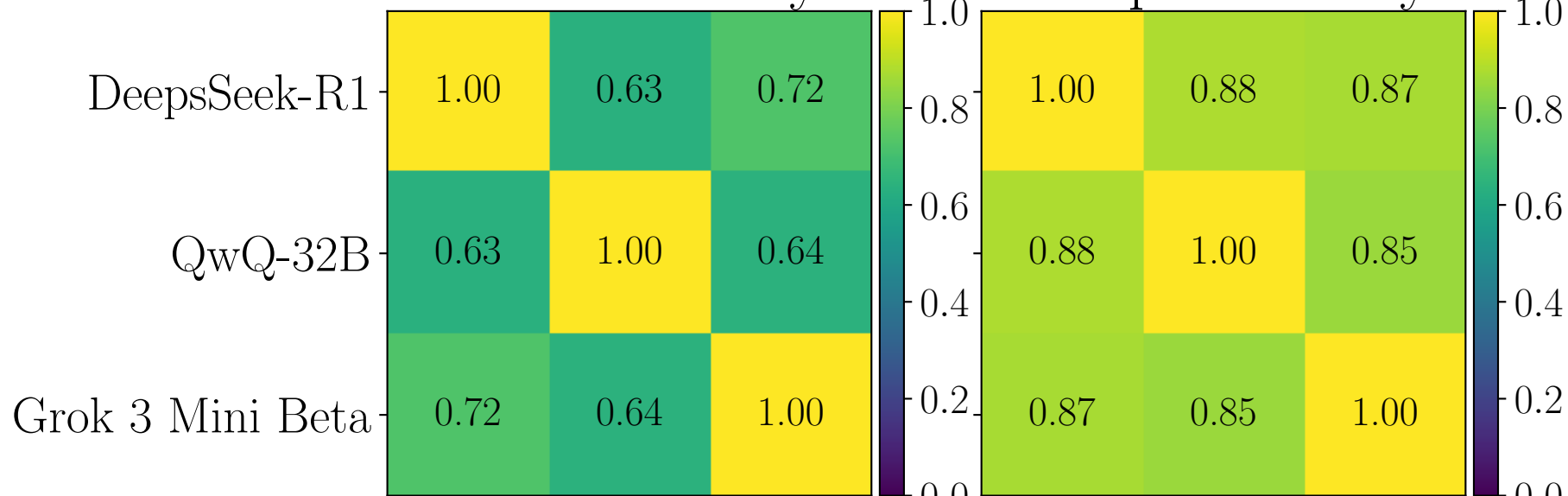


MATH-500

Tree Similarity

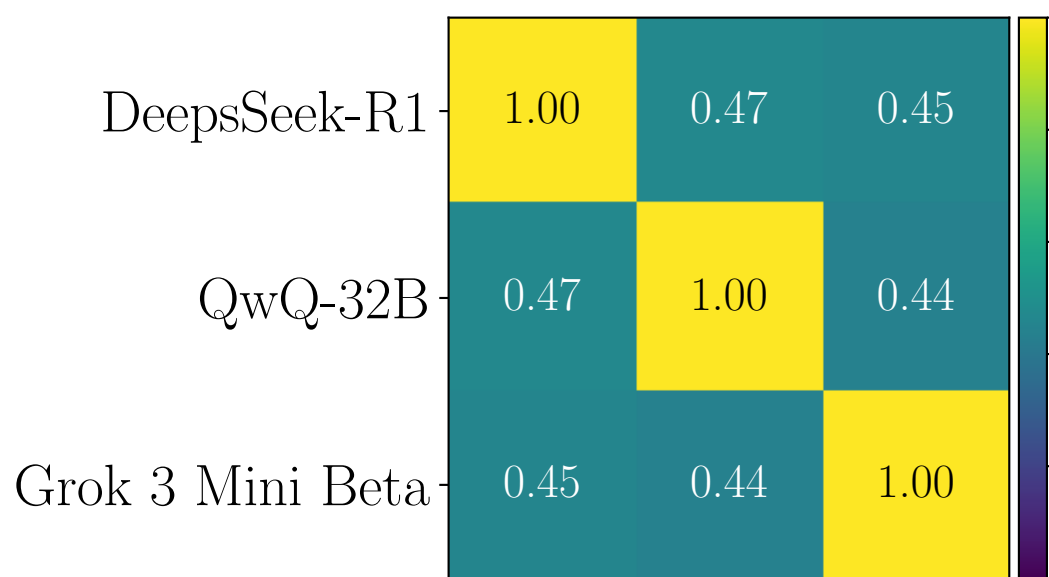


Jump Similarity

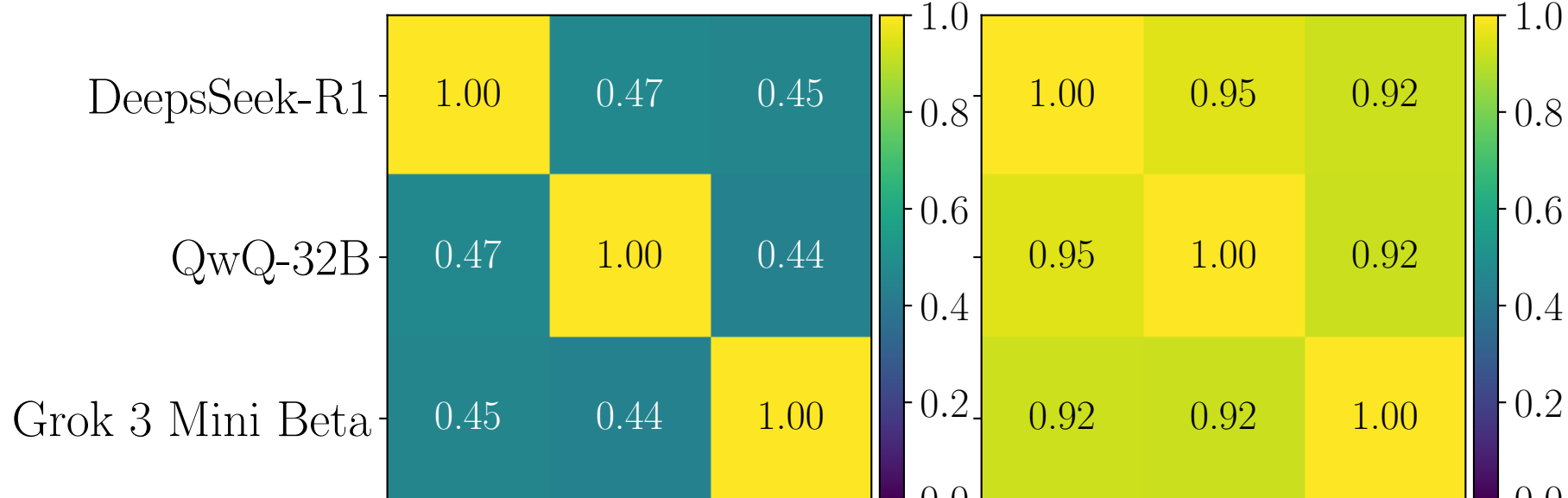


Game of 24

Tree Similarity



Jump Similarity



DeepSeek-R1

QwQ-32B

Grok 3 Mini Beta

DeepSeek-R1

QwQ-32B

Grok 3 Mini Beta