

Multi-Dimensional Comparison: Accuracy Across Error Categories

- Zero-Shot
- Chain-of-Thought
- LogiCue

Modal Negation

Accuracy

Probability Preservation

Defeasibility

100%

75%

50%

25%

