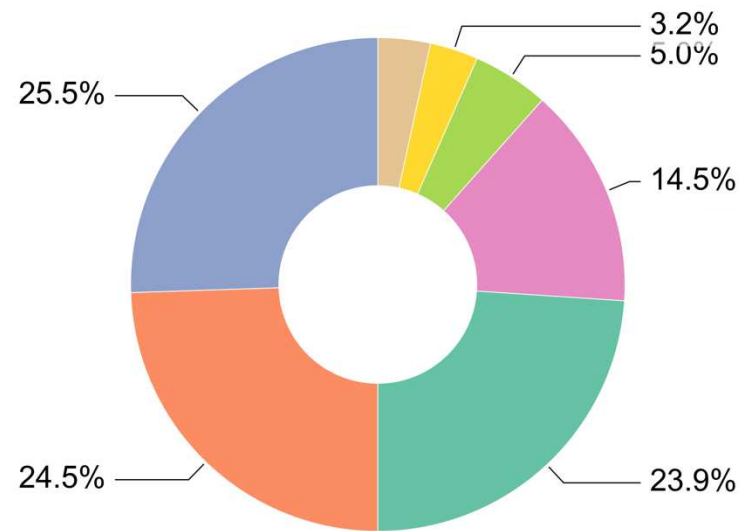
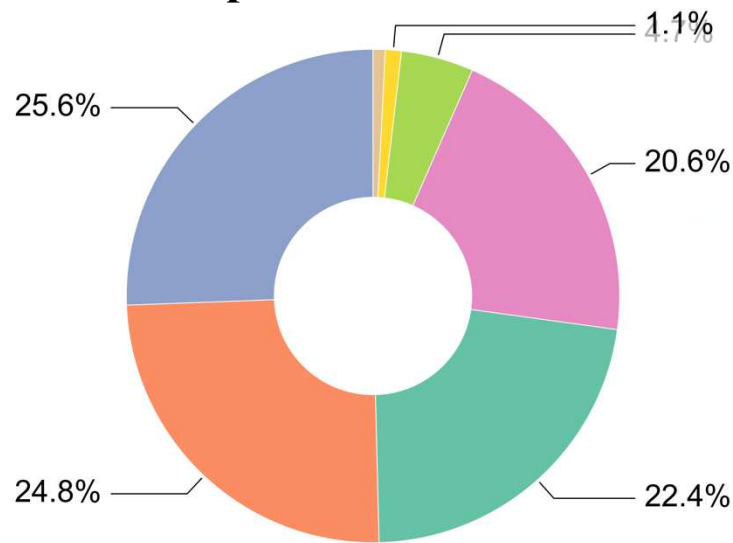


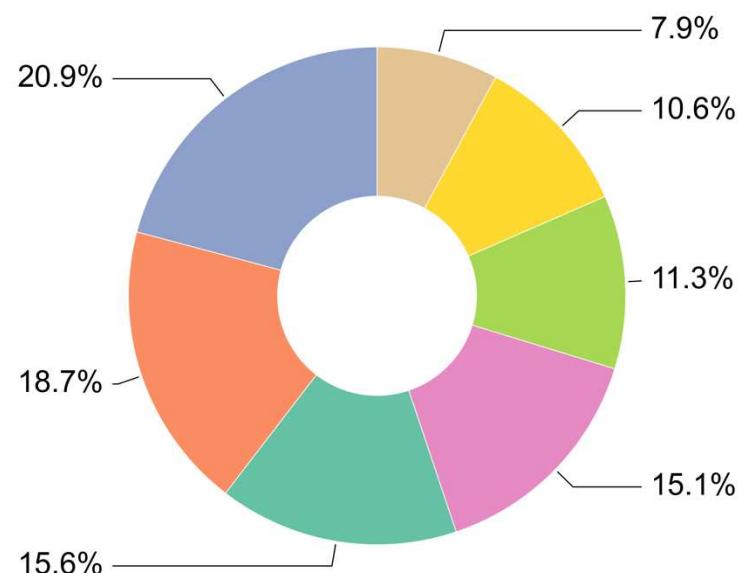
### Claude-3.7-Sonnet-Thinking



### OpenAI-GPT4.1



### Qwen2.5-VL-Instruct-70B



Invalid/unproven proof steps

Missing preconditions/hidden assumptions

Missing multimodal information

Formalization-natural language mismatch

Incorrect/undeclared variables/definitions

Others

Hallucinated/unjustified constructs