

9. Window Sizing

Working the budget backwards

An AI agent must support a 30-turn conversation before its window fills. Each turn adds 250 tokens, and the system prompt is 500 tokens. What window does the agent need?

Step 1: Tokens needed for all the turns

$$30 \times 250 = 7,500 \text{ tokens}$$

Step 2: Add the system prompt

$$7,500 + 500 = 8,000 \text{ tokens}$$

Step 3: Take the smallest window that fits

☐ 4k ☒ 8k ☐ 16k ☐ 32k

Practice 1

An AI agent must support a 40-turn conversation before its window fills. Each turn adds 300 tokens, and the system prompt is 800 tokens. What window does the agent need?

Step 1: Tokens needed for all the turns

Step 2: Add the system prompt

Step 3: Take the smallest window that fits

☐ 4k ☐ 8k ☐ 16k ☐ 32k

9. Window Sizing (continued)

Practice 2

An AI agent must support a 10-turn conversation before its window fills. Each turn adds 300 tokens, and the system prompt is 500 tokens. What window does the agent need?

Step 1: Tokens needed for all the turns

Step 2: Add the system prompt

Step 3: Take the smallest window that fits

☐ 4k
 ☐ 8k
 ☐ 16k
 ☐ 32k