2. Window Occupancy

System prompt, user, model

An AI agent has a 4,000-token context window. It holds a 400-token system prompt, 1,400 user tokens, and 1,000 output tokens. Mark the system prompt S, the user tokens U, and the output tokens O, then work out the free space.

Step 1: Mark what each part holds

1 - 1,000	S	S	S	S	U	U	U	U	U	U
1,001 - 2,000	U	U	U	U	U	U	U	U	O	O
2,001 - 3,000	O	O	O	O	O	O	O	O		
3,001 - 4,000										

Step 2: Subtract the used tokens from the window

4,000 - 400 - 1,400 - 1,000 = 1,200 tokens free

Practice 1

An AI agent has a 4,000-token context window. It holds a 300-token system prompt, 1,100 user tokens, and 800 output tokens. Mark the system prompt S, the user tokens U, and the output tokens O, then work out the free space.

Step 1: Mark what each part holds

1 - 1,000										
1,001 - 2,000										
2,001 - 3,000										
3,001 - 4,000										

Step 2: Subtract the used tokens from the window

2. Window Occupancy (continued)

Practice 2

An AI agent has a 4,000-token context window. It holds a 200-token system prompt, 1,000 user tokens, and 800 output tokens. Mark the system prompt S, the user tokens U, and the output tokens O, then work out the free space.

Step 1: Mark what each part holds

1 - 1,000									
1,001 - 2,000									
2,001 - 3,000									
3,001 - 4,000									

Step 2: Subtract the used tokens from the window