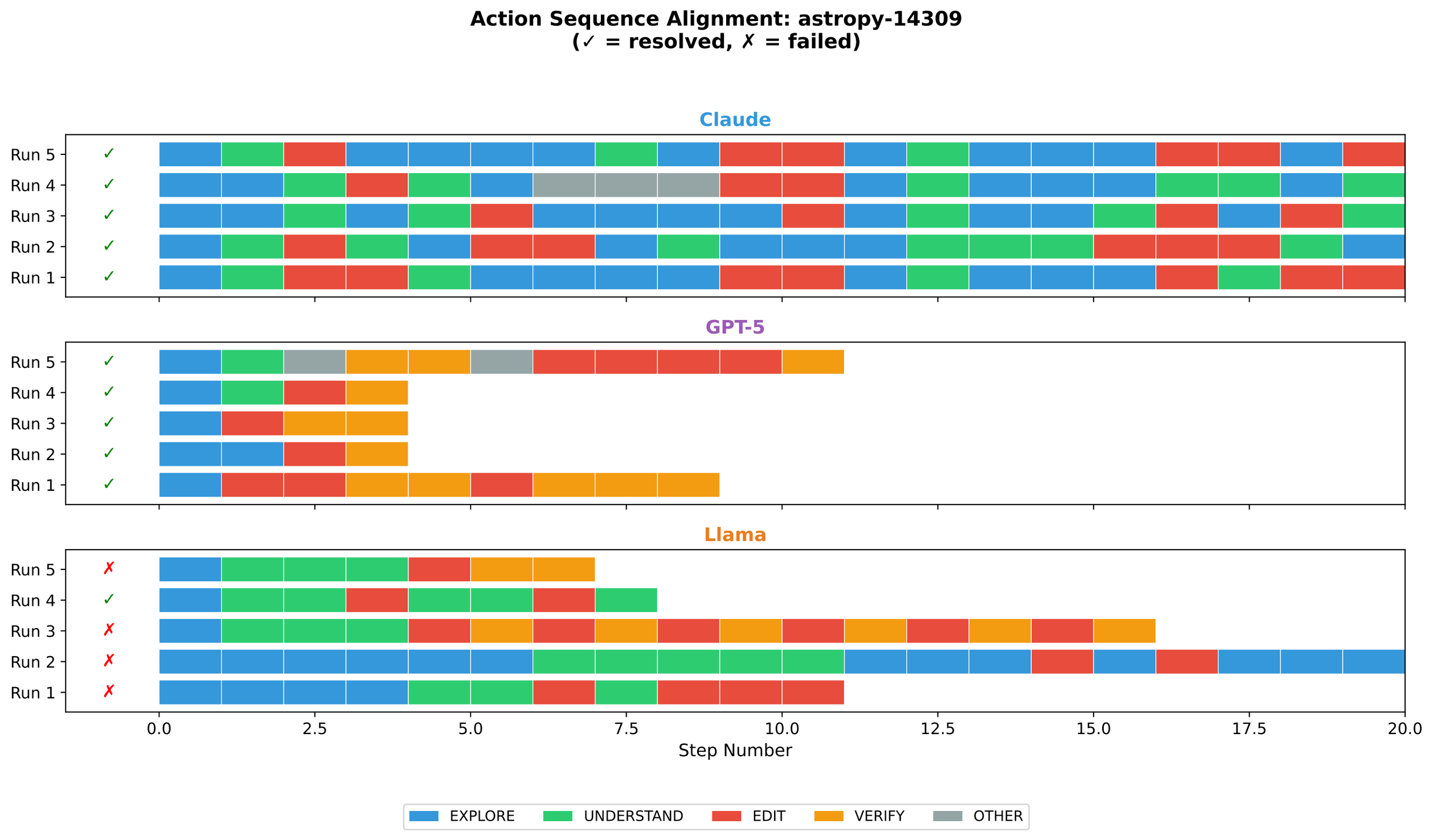
Action Sequence Alignment: astropy-14309
(✓ = resolved, ✗ = failed)
Claude
Run 5
Run 4
Run 3
Run 2
Run 1
GPT-5
Llama
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
Step Number
EXPLORE
UNDERSTAND
EDIT
VERIFY
OTHER