Human vs o3-mini: Compiled Successes Among Baseline Files

425

805

111

o3-mini
(916/1515, 60.5%)

Human
(1230/1515, 81.2%)