gemini-1.5-flash, BostonMarathon 0.7 0.6 0.2 0.1 0.0 DetConstSort BASK REP EXP