Sokoban R^P R^P R^P R^P R^P Alignment0.01 Alignment100 Performance0.01 Performance100 40 -40 -0.4 -4000 --250 **-**30 -30 -0.3 -3000 -20 -20 **-**0.2 -2000 --500 **-**10-10 -1000 0.1 -–750 **-**0 -0.0 --10 **-**-0.1 --1000 6000 2000 6000 6000 2000 4000 2000 6000 6000 R^A R^A R^A R^A R^A Alignment0.01 Alignment100 Performance0.01 Performance100 0.00 -0 --0.985 **-**-0.985 **-**-0.00985 **-**-0.25 **-**-25 **-**-0.990-0.990 **-**ROLF -0.00990 **-**Score -50 **-**-0.50 **-**-0.995 **-**-0.995-0.00995 **-**-75 **-**-0.75 --1.000 --0.01000 **-**-100 -1.00 -1.000 **-**6000 4000 4000 4000 4000 4000 2000 2000 2000 2000 6000 6000 2000 6000 6000 R^* R^* R^* R^* R^* Alignment0.01 Alignment100 Performance0.01 Performance100 40 --10 **-**-10 --10 **-**20 --250 **-**-20 **-**-20 **-**-20 **-**0 -SEBA -30 **-**-30 **-**-30 **-**-500 **-**-20 **-**-40 **--40** -40-750 **-**-50 **-**-50 **-**4000 6000 2000 4000 6000 2000 6000 4000 6000 2000 4000 2000 6000 2000 4000 Episode number

