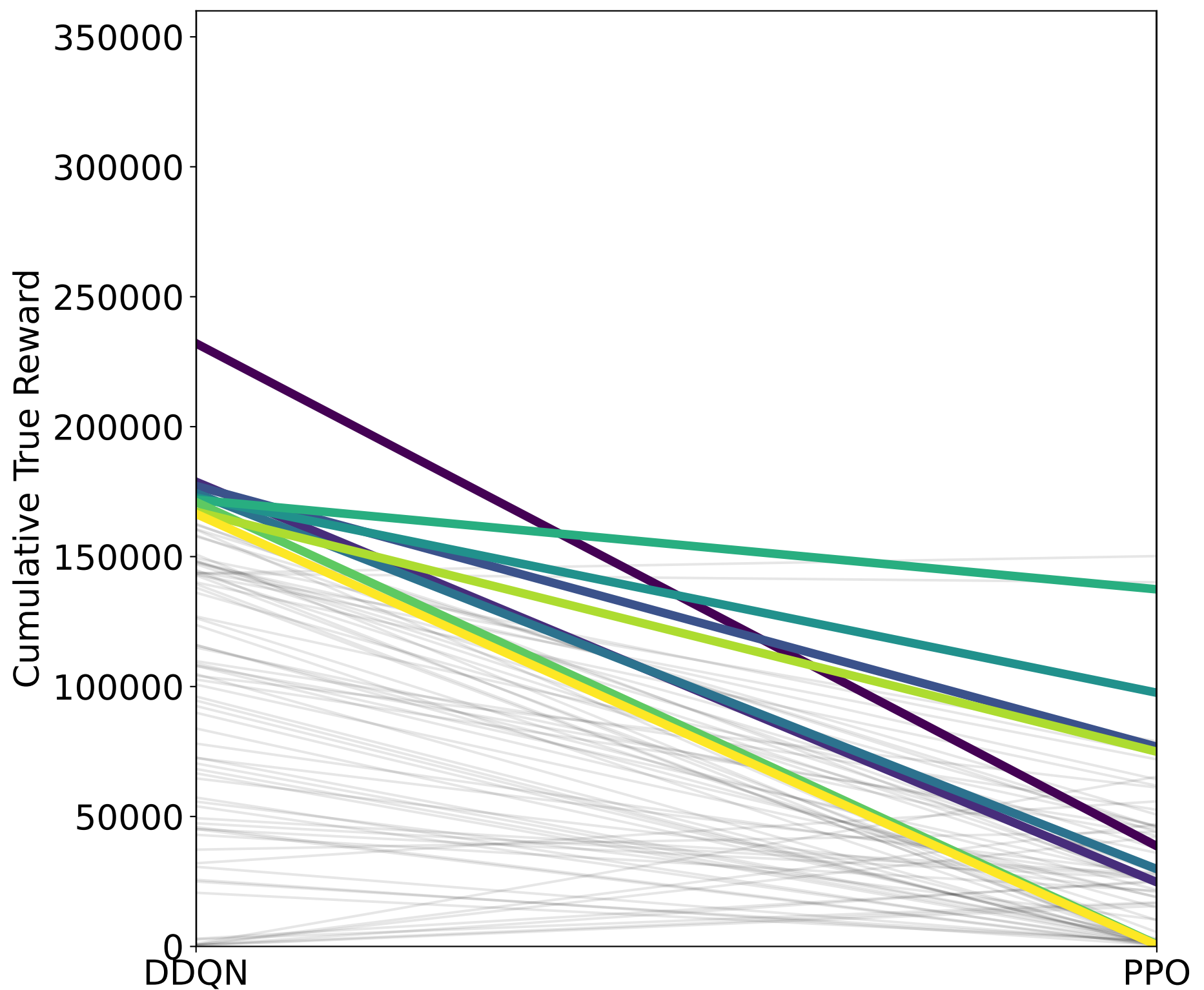


Relative Reward Function Performance  
for DDQN and PPO



	H $\wedge$ T:	-0.10;	H $\wedge$ $\neg$ T:	0.20;	$\neg$ H $\wedge$ T:	0.50,	$\neg$ H $\wedge$ $\neg$ T:	1.00
	H $\wedge$ T:	-0.40;	H $\wedge$ $\neg$ T:	-0.50;	$\neg$ H $\wedge$ T:	0.00,	$\neg$ H $\wedge$ $\neg$ T:	1.00
	H $\wedge$ T:	-1.00;	H $\wedge$ $\neg$ T:	-1.00;	$\neg$ H $\wedge$ T:	10.00,	$\neg$ H $\wedge$ $\neg$ T:	10.00
	H $\wedge$ T:	-1.00;	H $\wedge$ $\neg$ T:	-0.60;	$\neg$ H $\wedge$ T:	-0.80,	$\neg$ H $\wedge$ $\neg$ T:	1.00
	H $\wedge$ T:	-0.50;	H $\wedge$ $\neg$ T:	-0.50;	$\neg$ H $\wedge$ T:	10.00,	$\neg$ H $\wedge$ $\neg$ T:	10.00
	H $\wedge$ T:	0.00;	H $\wedge$ $\neg$ T:	0.00;	$\neg$ H $\wedge$ T:	5.50,	$\neg$ H $\wedge$ $\neg$ T:	10.00
	H $\wedge$ T:	-1.00;	H $\wedge$ $\neg$ T:	-0.70;	$\neg$ H $\wedge$ T:	-0.50,	$\neg$ H $\wedge$ $\neg$ T:	1.00
	H $\wedge$ T:	-0.10;	H $\wedge$ $\neg$ T:	-0.10;	$\neg$ H $\wedge$ T:	-0.10,	$\neg$ H $\wedge$ $\neg$ T:	1.00
	H $\wedge$ T:	-0.50;	H $\wedge$ $\neg$ T:	0.10;	$\neg$ H $\wedge$ T:	0.10,	$\neg$ H $\wedge$ $\neg$ T:	1.00