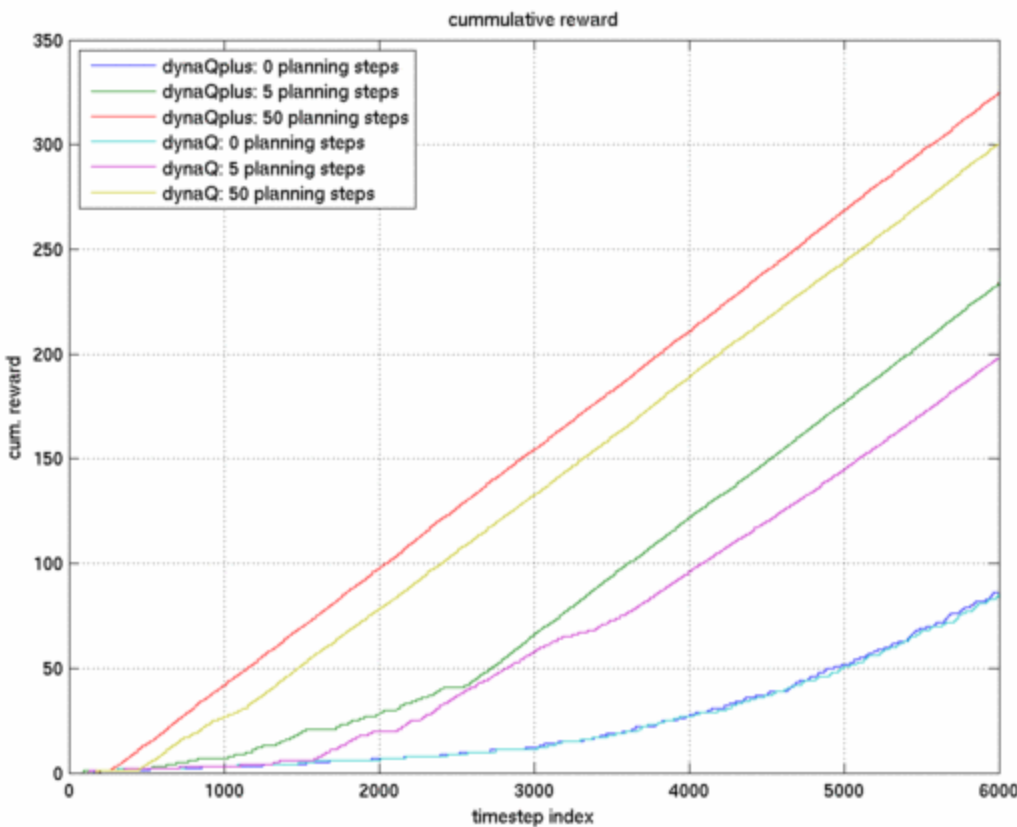


Here you will find experiments and results obtained when performing various different planning lengths in the dynaQ and dynaQplus algorithm for the "shortcut" maze, where at the 3000th timestep a shortcut opens up that would allow an agent to get more reward with fewer timesteps. The planning lengths chosen were 0 (no planning), 5, and 50. When planning 50 iterations ahead it seems to be very important to specify the initial condition on the action value function correctly. If this is specified incorrectly convergence may be difficult. Learning curves with various amounts of planning are plotted here. In all cases we see the benefit that dynaQplus provides is apparent since it obtains more total reward.



This result looks quite similar to that presented in the book.

[John Weatherwax](#)

Last modified: Sun May 15 08:46:34 EDT 2005