WFDR for a1true-a2false-b1false-c1false, reward function 1