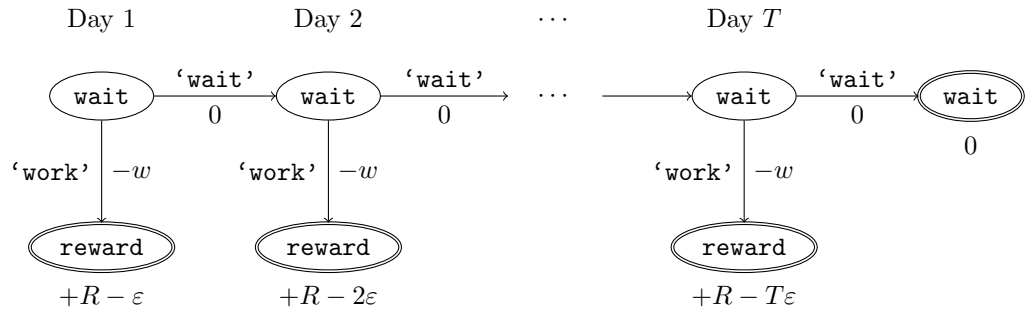
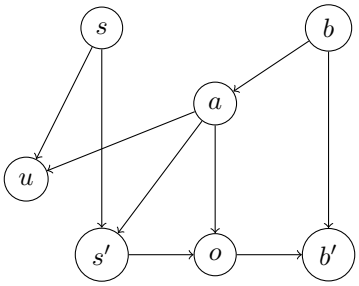
1 Procrastination MDP

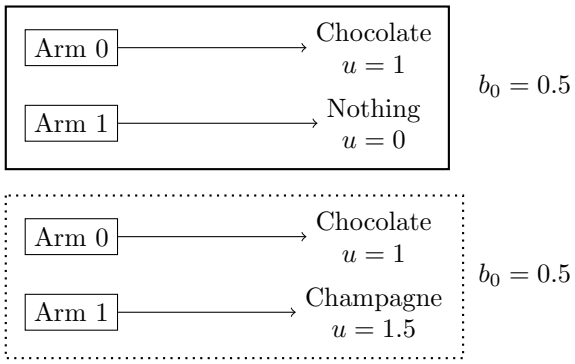


2 POMDP model

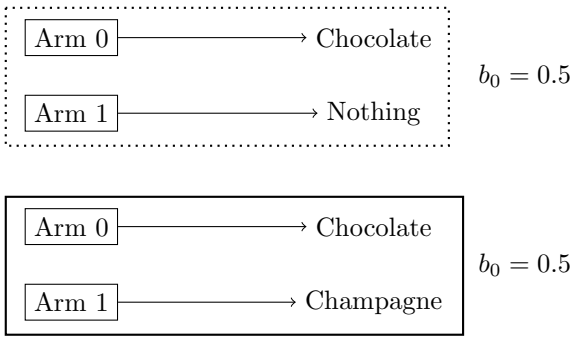


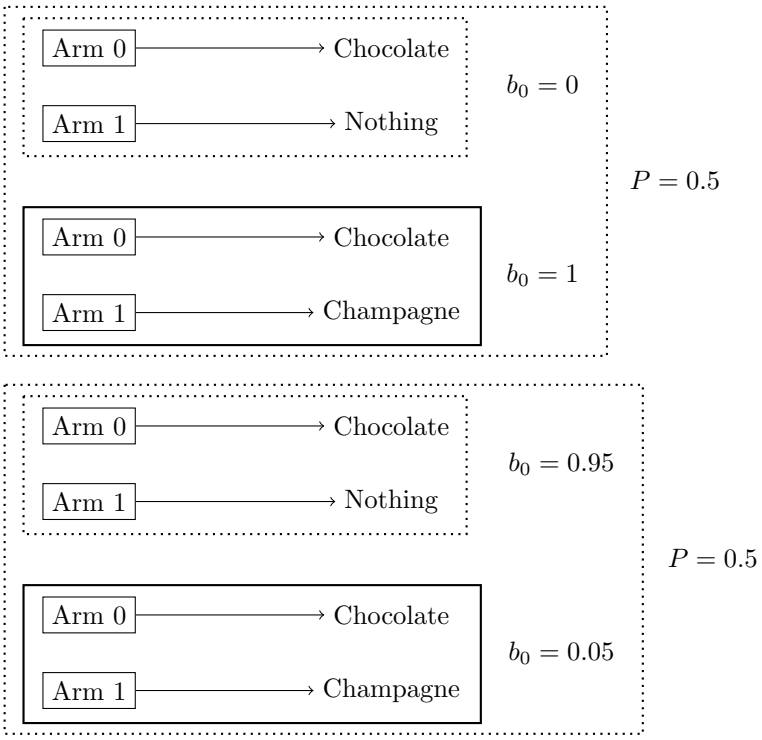
3 IRL Bandits

3.1 3c, first todo

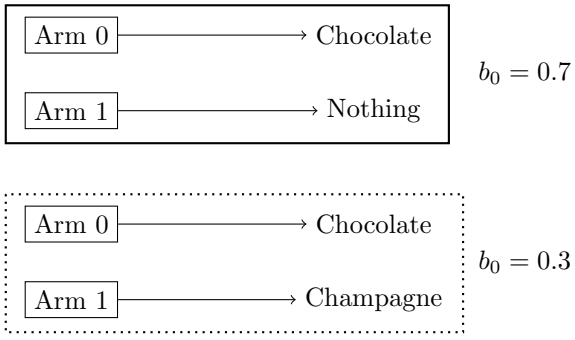


3.2 4, near end



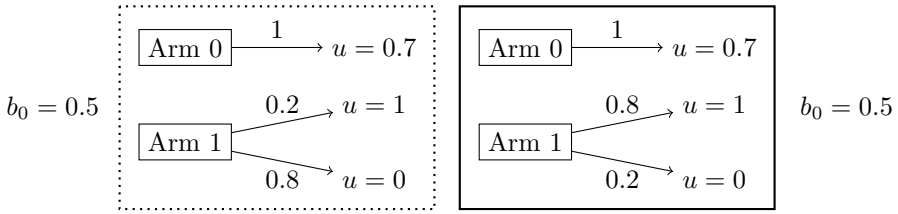

Arm 0
Chocolate
Arm 1
Nothing
$b_0 = 0$
Arm 0
Chocolate
Arm 1
Champagne
$b_0 = 1$
$P = 0.5$
Arm 0
Chocolate
Arm 1
Nothing
$b_0 = 0.95$
Arm 0
Chocolate
Arm 1
Champagne
$b_0 = 0.05$
$P = 0.5$

3.3 5c

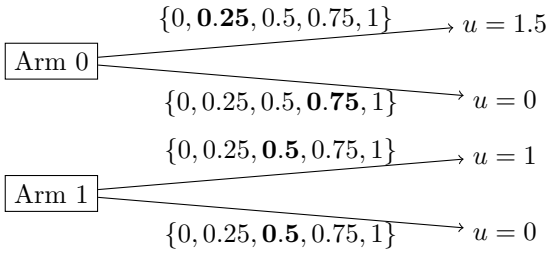

Arm 0
Chocolate
Arm 1
Nothing
$b_0 = 0.7$
Arm 0
Chocolate
Arm 1
Champagne
$b_0 = 0.3$

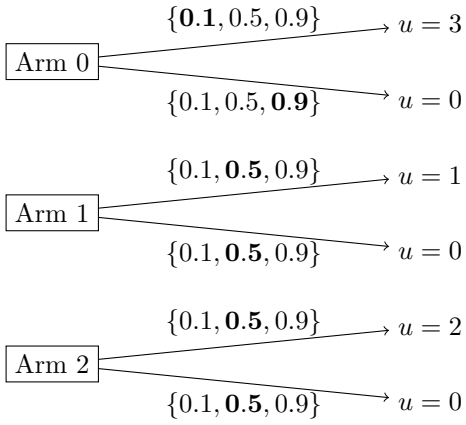
4 Stochastic bandits

4.1 3c

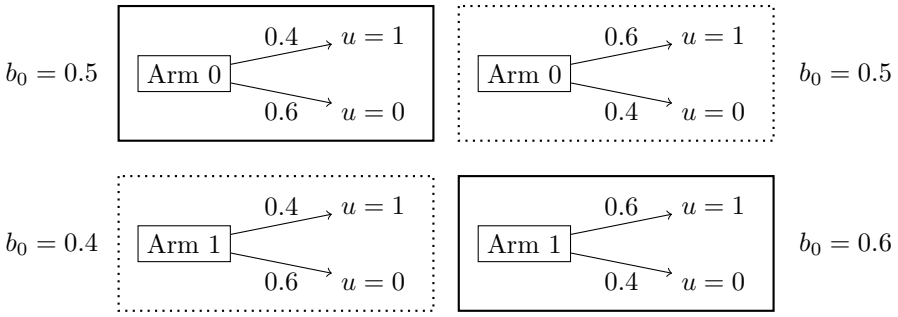


4.2 5b greedy





4.3 5b myopic


$b_0 = 0.5$
Arm 0
0.4
$u = 1$
0.6
$u = 0$
Arm 0
0.6
$u = 1$
0.4
$u = 0$
$b_0 = 0.5$
$b_0 = 0.4$
Arm 1
0.4
$u = 1$
0.6
$u = 0$
Arm 1
0.6
$u = 1$
0.4
$u = 0$
$b_0 = 0.6$