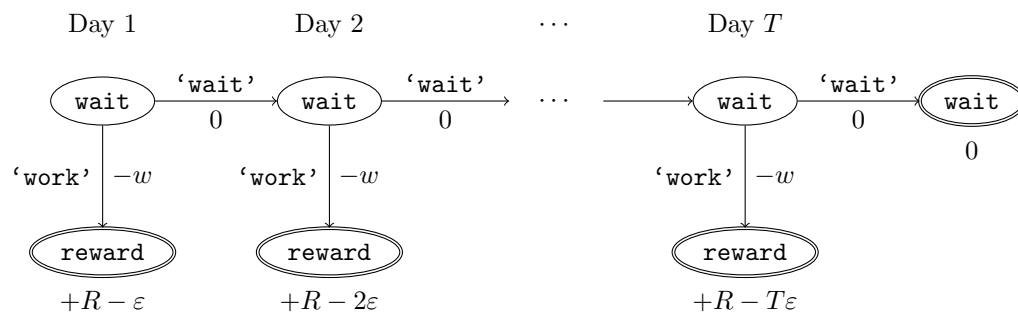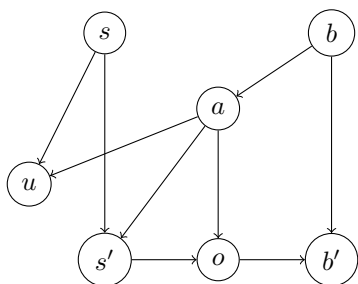# 1 Procrastination MDP

2 POMDP model

3 IRL Bandits

3.1 3c, first todo

Arm 0 ──────→ Chocolate $u = 1$	
Arm 1 ──────→ Nothing $u = 0$	$b_0 = 0.5$

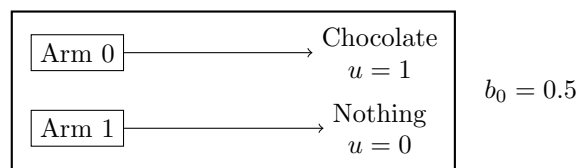|	
Arm 0 ──────→ Chocolate $u = 1$	
Arm 1 ──────→ Champagne $u = 1.5$	$b_0 = 0.5$

3.2 4, near end

Arm 0 ──────→ Chocolate	
Arm 1 ──────→ Nothing	$b_0 = 0.5$

Arm 0 ──────→ Chocolate	
Arm 1 ──────→ Champagne	$b_0 = 0.5$

3.3 5a

3.4   5c

Arm 0 ──────→ Chocolate

Arm 1 ──────→ Nothing

$b_0 = 0.7$

Arm 0 ──────→ Chocolate

Arm 1 ──────→ Champagne

$b_0 = 0.3$
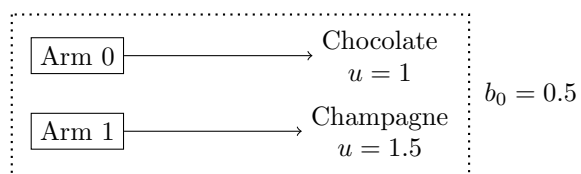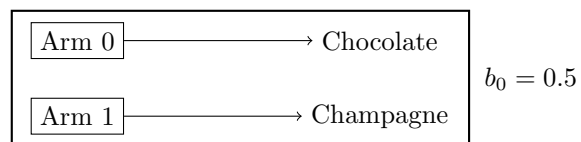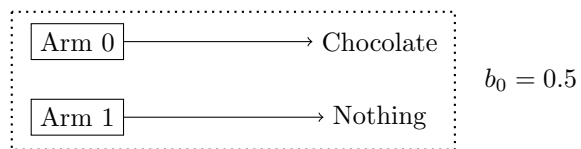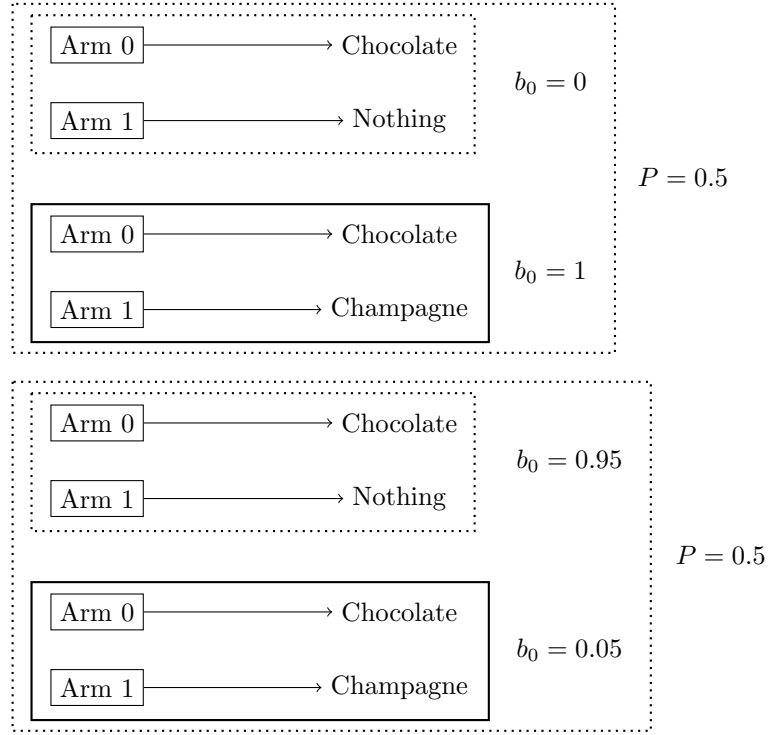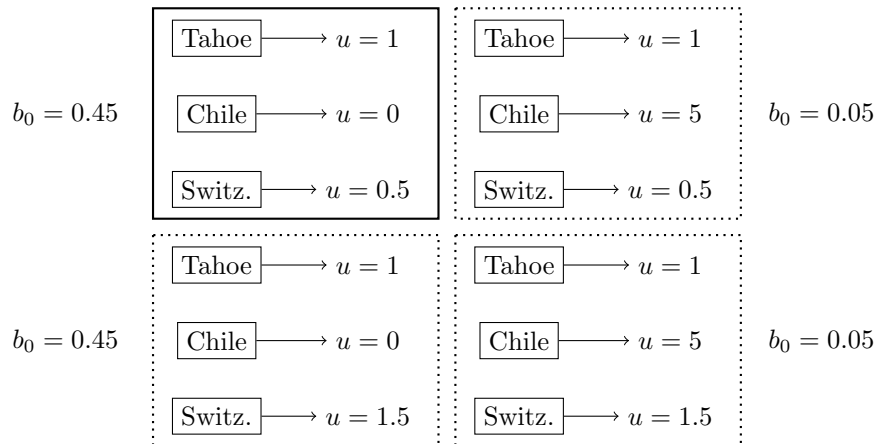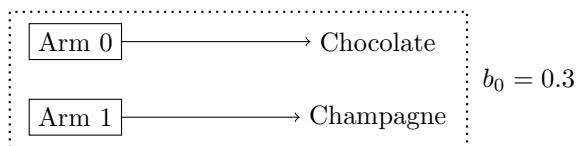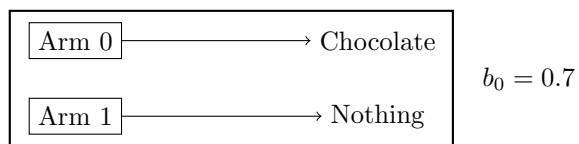
4   Stochastic bandits

4.1   3c

$b_0 = 0.5$

Arm 0 ──1──→ $u = 0.7$

Arm 1 ──0.2──→ $u = 1$

Arm 1 ──0.8──→ $u = 0$

Arm 0 ──1──→ $u = 0.7$

Arm 1 ──0.8──→ $u = 1$

Arm 1 ──0.2──→ $u = 0$

$b_0 = 0.5$

4.2 5b greedy

$\{0, \mathbf{0.25}, 0.5, 0.75, 1\} \longrightarrow u = 1.5$

Arm 0

$\{0, 0.25, 0.5, \mathbf{0.75}, 1\} \longrightarrow u = 0$

$\{0, 0.25, \mathbf{0.5}, 0.75, 1\} \longrightarrow u = 1$

Arm 1

$\{0, 0.25, \mathbf{0.5}, 0.75, 1\} \longrightarrow u = 0$

$\{\mathbf{0.1}, 0.5, 0.9\} \longrightarrow u = 3$

Arm 0

$\{0.1, 0.5, \mathbf{0.9}\} \longrightarrow u = 0$

$\{0.1, \mathbf{0.5}, 0.9\} \longrightarrow u = 1$

Arm 1

$\{0.1, \mathbf{0.5}, 0.9\} \longrightarrow u = 0$

$\{0.1, \mathbf{0.5}, 0.9\} \longrightarrow u = 2$

Arm 2

$\{0.1, \mathbf{0.5}, 0.9\} \longrightarrow u = 0$

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 → $u = 0$
$b_0 = 0.6$