

Game Theory: Homework

Caio Figueiredo

December 4, 2019

1 Question 1

1.1 (a)

We can rewrite $V(\mu)$ as:

$$\begin{aligned} V(\mu) &= (1 - \delta)V^*(\mu) \\ V^*(\mu) &= \max_{a \in \{0,1\}} \{(\mathbb{E}(y_t) - \frac{1}{2})a + \delta\mathbb{E}(V^*(\mu'))\} \\ &= \max_{a \in \{0,1\}} \{(\frac{1}{2}\mu - \frac{1}{4})a + \delta\mathbb{E}(V^*(\mu'))\} \end{aligned} \tag{1}$$

where:

$$\begin{aligned} \mu' &= \mu \text{ if } a = 0 \\ \mu' &= \frac{\frac{3}{4}\mu}{\frac{1}{4} + \frac{1}{2}\mu} \text{ if } a = 1, y = 1 \\ \mu' &= \frac{\frac{1}{4}\mu}{\frac{3}{4} - \frac{1}{2}\mu} \text{ if } a = 1, y = 0 \end{aligned}$$

Now we can use the standard Recursive Dynamic proof, define:

$$T(f)(x) = \max_{a \in \{0,1\}} \{(\frac{1}{2}\mu - \frac{1}{4})a + \delta f(\mu')\}$$

T maps weakly increasing functions in weakly increasing functions because μ' is strictly increasing in μ and $(\frac{1}{2}\mu - \frac{1}{4})a$ is strictly increasing as long as $a = 1$ and constant as long as $a = 0$.

Therefore by the Contraction Mapping Theorem the fix point of T , which is V^* , is weakly increasing. Moreover, since V^* is weakly increasing and $(1 - \delta) > 0$, then V is weakly increasing.

1.2 (b)

First notice that we can use the same argument of (a) to show that V^* is continuous.

Now notice that for $\mu = 0$, $\mu' = 0 \forall a, y$ and for $\mu = 1$, $\mu' = 1 \forall a, y$, therefore:

$$\begin{aligned} V^*(0) &= \max_{a \in \{0,1\}} \{-\frac{1}{4}a + \delta V^*(0)\} & V^*(1) &= \max_{a \in \{0,1\}} \{\frac{1}{4}a + \delta V^*(1)\} \\ &= \frac{\max_{a \in \{0,1\}} \{-\frac{1}{4}a\}}{(1 - \delta)} & &= \frac{\max_{a \in \{0,1\}} \{\frac{1}{4}a\}}{(1 - \delta)} \\ &= 0 & &= \frac{1}{4(1 - \delta)} > 0 \\ V(0) &= 0 & V(1) &= \frac{1}{4} \end{aligned}$$

Since V is weakly increasing that must be $\mu^* \geq 0$ such that for for all $0 \leq \mu < \mu^*$, $V(\mu) = 0$, and for all $1 \geq \mu > \mu^*$, $V(\mu) > 0$.

Moreover, $0 \leq \mu < \mu^*$, optimal a is 0 and for $\mu^* < \mu \leq 1$, optimal a is 1. Since V^* is continuous at μ^* the agent is indifferent.

1.3 (c)

First notice that the statement is not true for $\mu = 1$, in this case the agent never update his belief and always choose $a = 1$.

However, for any other μ , such that $\mu^* < \mu < 1$, there exists finite number of bad signals in sequence that is enough to bring it down to a number below μ^* . To see that define: μ_0^n the belief after observing $y = 0, n$ times in a row. It easy to show that:

$$\mu_0^n = \frac{\mu}{3^n - \sum_{i=0}^n (3^{i-1} 2) \mu}$$

It is also simple to show that $\mu_0^n \rightarrow 0$ as $n \rightarrow \infty$. Therefore: we have for any $\epsilon > 0$, in particular, for $\epsilon = \mu^*$, $\exists n^* \in \mathbb{N}$ such that for any $n > n^*$, $\mu_0^n < \mu^*$.

We already establish that once below μ^* the agent always choose $a = 0$, which implies he don't update his belief anymore and therefore always choose $a = 0$. Since n^* is a finite number the probability of observing the bad signal n^* times is positive and this part of the statement is proven.

We are left to prove that there is a probability that the belief never falls below μ^* .

I now define a new state:

$$s_t = (\# \text{ Good Signals}) - (\# \text{ Bad Signals})$$

Also define $P_\infty(s)$ as the probability that at state s the agent only plays $a = 1$ forever. Notice that by the same argument as above, as $s \rightarrow \infty$ the belief goes to 1 and therefore $P_\infty(\infty) = 1$. We have:

$$P_\infty(s) = P(y = 1)P_\infty(s + 1) + P(y = 0)P_\infty(s - 1)$$

Which follows from the definition of P_∞ . Moreover since $\theta = 1 : P(y = 1) = \frac{3}{4}$ and $P(y = 0) = \frac{1}{4}$. For simplification we further assume that the prior μ is such that at $s = 0$ one bad return is enough to bring our belief below μ^* , which implies that $P_\infty(-1) = 0$. This is WLOG because if the affirmation hold for a μ this low is must also holds for higher μ . Therefore:

$$\begin{aligned} P_\infty(0) &= \frac{3}{4}P_\infty(1) \\ &= \frac{3}{4}\left(\frac{3}{4}P_\infty(2) + \frac{1}{4}P_\infty(0)\right) \\ \frac{13}{16}P_\infty(0) &= \frac{9}{16}P_\infty(2) \end{aligned}$$

Therefore:

$$\begin{aligned} P_\infty(1) &= \frac{4}{3}P_\infty(0) \\ P_\infty(2) &= \frac{13}{9}P_\infty(0) \end{aligned}$$

Moreover:

$$\begin{aligned} P_\infty(n) &= \frac{3}{4}P_\infty(n + 1) + \frac{1}{4}P_\infty(n - 1) \\ \frac{3}{4}P_\infty(n + 1) &= P_\infty(n) - \frac{1}{4}P_\infty(n - 1) \\ P_\infty(n + 1) &= \frac{4}{3}P_\infty(n) - \frac{1}{3}P_\infty(n - 1) \\ P_\infty(n + 1) - \frac{1}{3}P_\infty(n) &= P_\infty(n) - \frac{1}{3}P_\infty(n - 1) \\ &= P_\infty(2) - \frac{1}{3}P_\infty(1) \\ &= \frac{13}{9}P_\infty(0) - \frac{1}{3}\frac{4}{3}P_\infty(0) \\ P_\infty(0) &= P_\infty(n + 1) - \frac{1}{3}P_\infty(n) \end{aligned}$$

As $n \rightarrow \infty$:

$$\begin{aligned} P_\infty(0) &= \frac{2}{3}P_\infty(\infty) \\ P_\infty(0) &= \frac{2}{3} \end{aligned}$$

Therefore the probability of playing $a = 1$ forever when $s = 0$, that is the initial state, is $\frac{2}{3}$ which concludes our proof.

1.4 (d)

We can use the very same technique as before. Redefine $P_\infty(s)$ as the probability that the agent eventually takes action 0 forever. Remember, $P_\infty(-\infty) = 1$.

$$\begin{aligned} P_\infty(n) &= \frac{1}{4}P_\infty(n+1) + \frac{3}{4}P_\infty(n-1) \\ P_\infty(n+1) &= 4P_\infty(n) - 3P_\infty(n-1) \\ P_\infty(n+1) - P_\infty(n) &= 3(P_\infty(n) - P_\infty(n-1)) \\ P_\infty(n+2) - P_\infty(n+1) &= 3^2(P_\infty(n) - P_\infty(n-1)) \\ P_\infty(\infty) - P_\infty(\infty) &= 3^\infty(P_\infty(n) - P_\infty(n-1)) \\ 0 &= P_\infty(n) - P_\infty(n-1) \\ P_\infty(n) &= P_\infty(n-1) \\ P_\infty(0) &= P_\infty(-1) = \dots = P_\infty(-\infty) = 1 \end{aligned}$$

2 Question 2

$$\begin{aligned} |V^c(\delta, \theta) - V(\delta, \theta)| &= \left| \sum_{t=0}^{\infty} \delta^t a_t^c \left(\sum_{i=1}^{\infty} y_i P_{\theta^*}(y_i | y_1, \dots, y_t) \right) - \sum_{t=0}^{\infty} \delta^t a_t \left(\sum_{i=1}^{\infty} y_i P_{\theta^*}(y_i | y_1, \dots, y_t) \right) \right| \\ &\leq \left| \sum_{t=0}^{\infty} \delta^t a_t^c \left(\sum_{i=1}^{\infty} y_i P_{\theta^*}(y_i | y_1, \dots, y_t) \right) - \sum_{t=0}^{\infty} \delta^t a_t \left(\sum_{i=1}^{\infty} y_i \mu(y_i | y_1, \dots, y_t) \right) \right| + \\ &\quad \left| \sum_{t=0}^{\infty} \delta^t a_t \left(\sum_{i=1}^{\infty} y_i P_{\theta^*}(y_i | y_1, \dots, y_t) \right) - \sum_{t=0}^{\infty} \delta^t a_t \left(\sum_{i=1}^{\infty} y_i \mu(y_i | y_1, \dots, y_t) \right) \right| \\ &\leq 2 \left(\sum_{t=0}^{\infty} \delta^{2t} \max_{y \in Y} |y|^2 \left(\sum_{i=1}^{\infty} |P_{\theta^*}(y_i | y_1, \dots, y_t) \mu(y_i | y_1, \dots, y_t)| \right)^2 \right)^{1/2} + \\ &\leq 2 \max_{y \in Y} |y| \left(\sum_{t=0}^{\infty} \delta^{2t} \right)^{1/2} \left(\sum_{t=0}^{\infty} \sup_{y \in Y} \{P_{\theta^*}(y | y_1, \dots, y_t) \mu(y | y_1, \dots, y_t)\}^2 |y| \right)^{1/2} \\ &\leq \frac{2|y|^{1/2} \max_{y \in Y} |y|}{(1 - \delta^2)^{1/2}} \left(\sum_{t=0}^{\infty} \frac{1}{2} KL(\theta^* | \mu) \right)^{1/2} \\ &\leq \frac{2|y|^{1/2} \max_{y \in Y} |y|}{(1 - \delta^2)^{1/2}} \left(\sum_{t=0}^{\infty} -\ln e^{-(1-\delta)^{2(t+1)}} \right)^{1/2} \\ &= \frac{2|y|^{1/2} \max_{y \in Y} |y|}{(1 - \delta^2)^{1/2}} \frac{1 - \delta}{(1 - (1 - \delta)^2)^{1/2}} \\ &= 2|y|^{1/2} \max_{y \in Y} |y| \left(\frac{1 - \delta}{(1 + \delta)(1 - (1 - \delta)^2)} \right)^{1/2} \end{aligned}$$

Finally, notice that $1 - \delta \rightarrow 0$ as $\delta \rightarrow 1$ while $(1 + \delta)(1 - (1 - \delta)^2) \rightarrow 2$. Therefore

$$\lim_{\delta \rightarrow 1} V^c(\delta, \theta) - V(\delta, \theta) = 0$$

3 Question 3

3.1 (a)

For simplification:

$$\begin{aligned}P_\gamma(y) &= \mu P_\gamma(y|\theta = 1) + (1 - \mu)P_\gamma(y|\theta = -1) \\P_\gamma(1) &= \mu(1/2 + \gamma) + (1 - \mu)(1/2\gamma) \\&= 1/2 + \gamma(2\mu - 1) \\P_\gamma(0) &= 1/2 - \gamma(2\mu - 1)\end{aligned}$$

Notice:

$$\lim_{\gamma \rightarrow 0} P_\gamma(0) = \lim_{\gamma \rightarrow 0} P_\gamma(1) = \frac{1}{2}$$

The Bayesian updates are:

$$\begin{aligned}\mu_\gamma(1) &= \frac{(1/2 + \gamma)\mu}{P_\gamma(1)} \\ \mu_\gamma(0) &= \frac{(1/2 - \gamma)\mu}{P_\gamma(0)}\end{aligned}$$

Therefore:

$$\lim_{\gamma \rightarrow 0} \mu_\gamma(0) = \lim_{\gamma \rightarrow 0} \mu_\gamma(1) = \mu$$

and:

$$\lim_{\gamma \rightarrow 0} V(\mu_\gamma(0)) = \lim_{\gamma \rightarrow 0} V(\mu_\gamma(1)) = V(\mu)$$

Finally:

$$\begin{aligned}V_\gamma(\mu) &= P_\gamma(0)V(\mu_\gamma(0)) + P_\gamma(1)V(\mu_\gamma(1)) \\&= (1/2 + \gamma(2\mu - 1))V(\mu_\gamma(1)) + (1/2 - \gamma(2\mu - 1))V(\mu_\gamma(0))\end{aligned}$$

and:

$$\lim_{\gamma \rightarrow 0} V_\gamma(\mu) = V(\mu)$$

Which make the question identity equals the derivative of V_γ in relation to γ :

$$\frac{\partial V_\gamma(\mu)}{\partial \gamma} = \lim_{\gamma \rightarrow 0} \frac{V_\gamma(\mu) - V(\mu)}{\gamma}$$

We are left to prove that this derivative is zero at $\gamma = 0$:

$$\begin{aligned}\frac{\partial V_\gamma(\mu)}{\partial \gamma}(\gamma) &= (2\mu - 1)V(\mu_\gamma(1)) + \frac{\partial V(\mu_\gamma(1))}{\partial \gamma}(1/2 + \gamma(2\mu - 1)) - \\&\quad (2\mu - 1)V(\mu_\gamma(0)) + \frac{\partial V(\mu_\gamma(0))}{\partial \gamma}(1/2 - \gamma(2\mu - 1)) \\ \frac{\partial V_\gamma(\mu)}{\partial \gamma}(0) &= (2\mu - 1)V(\mu) + \frac{\partial V(\mu_0(1))}{\partial \gamma}(1/2) - (2\mu - 1)V(\mu) + \frac{\partial V(\mu_0(0))}{\partial \gamma}(1/2) \\&= 1/2 \left(\frac{\partial V(\mu_0(0))}{\partial \gamma} + \frac{\partial V(\mu_0(1))}{\partial \gamma} \right)\end{aligned}$$

But:

$$\begin{aligned}
V(\mu_\gamma(1)) &= \mu_\gamma u(a_1^*, 1) + (1 - \mu_\gamma)u(a_1^*, -1) \\
&= P_\gamma(1)^{-1}[(1/2 + \gamma)\mu u(a_1^*, 1) + \gamma(2\mu - 2)\mu u(a_1^*, -1)] \\
\frac{\partial V(\mu_\gamma(1))}{\partial \gamma} &= P_\gamma(1)^{-1}[\mu u(a_1^*) + (2\mu - 2)u(a_1^*, -1)] + \\
&\quad \frac{1 - 2\mu}{(1/2 + \gamma(2\mu - 1))^2}[(1/2 + \gamma)\mu u(a_1^*, 1) + \gamma(2\mu - 2)u(a_1^*, -1)] \\
\frac{\partial V(\mu_0(1))}{\partial \gamma} &= 2[\mu u(a_1^*) + (2\mu - 2)u(a_1^*, -1)] + (4 - 8\mu)[1/2\mu u(a_1^*, 1)]
\end{aligned}$$

Where a_1^* is the argmax of V for $\gamma = 0$ and:

$$\begin{aligned}
V(\mu_\gamma(0)) &= P_\gamma(0)^{-1}[(1/2 - \gamma)\mu u(a_0^*, 1) - \gamma(2\mu - 2)\mu u(a_0^*, -1)] \\
\frac{\partial V(\mu_\gamma(0))}{\partial \gamma} &= P_\gamma(0)^{-1}[-\mu u(a_0^*) - (2\mu - 2)u(a_0^*, -1)] + \\
&\quad \frac{2\mu - 1}{(1/2 - \gamma(2\mu - 1))^2}[(1/2 - \gamma)\mu u(a_0^*, 1) - \gamma(2\mu - 2)u(a_0^*, -1)] \\
\frac{\partial V(\mu_0(0))}{\partial \gamma} &= 2[-\mu u(a_0^*) - (2\mu - 2)u(a_0^*, -1)] + (8\mu - 4)[1/2\mu u(a_0^*, 1)]
\end{aligned}$$

And since $a_1^* \rightarrow a_0^*$ as $\gamma \rightarrow 0$, by what we have seen before, we have:

$$\begin{aligned}
\frac{\partial V(\mu_0(0))}{\partial \gamma} + \frac{\partial V(\mu_0(1))}{\partial \gamma} &= 0 \\
\frac{\partial V_\gamma(\mu)}{\partial \gamma}(0) &= 0 \\
\lim_{\gamma \rightarrow 0} \frac{V_\gamma(\mu) - V(\mu)}{\gamma} &= 0
\end{aligned}$$

■

3.2 (b)

The agent choose γ to maximize:

$$\max_{\gamma} V_\gamma(\mu) - \beta\gamma \quad (2)$$

Notice that as long as $\frac{\partial V_\gamma}{\partial \gamma}(\mu) < \beta$ the agent has no incentive to invest in learning. Moreover, as $\mu \rightarrow 0, \gamma \rightarrow 0$ since at $\mu = 0$ you have no incentive to learn ($\gamma = 0$) and $\frac{\partial V_\gamma}{\partial \gamma}(\mu)$ is continuous and μ .

From question (a) we have that: $\frac{\partial V_\gamma}{\partial \gamma}(\mu) \rightarrow 0$ as $\gamma \rightarrow 0$, therefore: $\exists \bar{\mu}$ such that $\forall \mu \in [0, \bar{\mu}] \frac{\partial V_\gamma}{\partial \gamma}(\mu) < \beta$, which imply that the optimal choice of γ is 0.

It is also easy to see that if $\gamma = 0$ the agent doesn't update and then $\mu_t = \mu_0$ for all t . Which concludes our proof.

3.3 (c)

I assume that $1/4 < \mu_0 < 3/4$ and that the agent is still myopic. Define:

$$u(a, \theta) = \frac{3}{16}a\theta \quad (3)$$

Which is 0 as long as the agent choose no $a = 0$, positive as long as he chooses $a = 1$ and the state $\theta = 1$, negative otherwise. Therefore if $\mu > 1/2$ he chooses $a = 1$, but choose $a = 0$ otherwise.

His expected payoff is:

$$\begin{aligned}
E(u(a, \theta)|\mu) - \gamma^3 &= \\
&= P(y = 1|\mu, \gamma) \max\{\mathbb{E}(u(1, \theta)|\mu_\gamma(1)), 0\} + \\
&\quad P(y = 0|\mu, \gamma) \max\{\mathbb{E}(u(1, \theta)|\mu_\gamma(0)), 0\} \\
&= [\mu(1/2 + \gamma) + (1 - \mu)(1/2 - \gamma)] \frac{3}{16} \frac{\mu - 1/2 + \gamma}{1/2 + 2\gamma\mu - \gamma} - \gamma^3 \\
&= [1/2 + 2\gamma\mu - \gamma] \frac{3}{16} \frac{\mu - 1/2 + \gamma}{1/2 + 2\gamma\mu - \gamma} - \gamma^3 \\
&= \frac{3}{16} [\mu - 1/2 + \gamma] - \gamma^3
\end{aligned} \tag{4}$$

FOC is:

$$\begin{aligned}
3/16 - 3\gamma^2 &= 0 \\
\gamma^2 &= \frac{1}{16} \\
\gamma &= \frac{1}{4}
\end{aligned} \tag{5}$$

And therefore by Berk Theorem he learns the true state with probability 1.

4 Question 4

4.1 (a)

$$\begin{aligned}
\phi_{\theta', \theta}^t(y_1, \dots, y_t) &= \log \frac{\mu^t(\theta'|y_1, \dots, y_t)}{\mu^t(\theta|y_1, \dots, y_t)} \\
&= \log \frac{\mu(\theta') P'_\theta(y_1, \dots, y_t)}{\mu(\theta) P_\theta(y_1, \dots, y_t)}
\end{aligned} \tag{6}$$

$$\begin{aligned}
P_\theta(y_1, \dots, y_t) &= \pi_0(y_1) \prod_{\tau=1}^{t-1} P_\theta(y_{\tau+1}|y_\tau) \\
\phi_{\theta', \theta}^t(y_1, \dots, y_t) &= \log \frac{\mu(\theta')}{\mu(\theta)} + \sum_{\tau=1}^{t-1} \log \frac{P'_\theta(y_{\tau+1}|y_\tau)}{P_\theta(y_{\tau+1}|y_\tau)}
\end{aligned} \tag{7}$$

$$\sum_{\tau=1}^{t-1} \log \frac{P'_\theta(y_{\tau+1}|y_\tau)}{P_\theta(y_{\tau+1}|y_\tau)} = \sum_{y, y'} (t-1) \rho_{y, y'} \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)}$$

But since $\rho_{y, y'}(y_1, \dots, y_t) = \rho_{y, y'}(y'_1, \dots, y'_t)$ we have $\mu^t(\theta|y_1, \dots, y_t) = \mu^t(\theta|y'_1, \dots, y'_t) \forall \theta \in \Theta$.

4.2 (b)

$$\begin{aligned}
\rho_{y, y'}(y_1, \dots, y_t) &\xrightarrow{P} \theta^*(y, y') = \pi^*(y) P_{\theta^*}(y'|y) \\
\phi_{\theta', \theta}^t(y_1, \dots, y_t) &= \log \frac{\mu(\theta')}{\mu(\theta)} + \sum_{y, y'} (t-1) \rho_{y, y'} \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)} \\
&= \log \frac{\mu(\theta')}{\mu(\theta)} + \sum_y \sum_{y'} (t-1) \rho_{y, y'} \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)}
\end{aligned} \tag{8}$$

But:

$$\begin{aligned}
\sum_{y'} (t-1) \rho_{y, y'} \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)} &\rightarrow \sum_{y'} \pi^*(y) P_{\theta^*}(y'|y) \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)} \\
&= \pi^*(y) (KL(P_{\theta^*}(\cdot|y)|P_\theta(\cdot|y)) - KL(P_{\theta^*}(\cdot|y)|P_\theta(\cdot|y)))
\end{aligned} \tag{9}$$

Therefore:

$$\sum_y \sum_{y'} (t-1) P_{y,y'} \log \frac{P'_\theta(y'|y)}{P_\theta(y'|y)} \rightarrow \sum_y \pi^*(y) (KL(P_{\theta^*}(\cdot|y)|P_\theta(\cdot|y)) - \sum_y KL(P_\theta^*(\cdot|y)|P_\theta(\cdot|y))) \quad (10)$$

And so for $\theta \notin \Theta^*$ and $\theta' \in \Theta^*$ $\phi_{\theta',\theta}^t \rightarrow \infty$ which implies that $\mu(\Theta^*|y_1, \dots, y_t) \xrightarrow{P} 1$

5 Question 5

5.1 (a)

I assume that $y^1, y^2 \in \{0, 1\}$.

Even though the agent don't observe y^1 and y^2 directly they can be inferred from the overall payoff. Since $y^1, y^2 \in \{0, 1\}$, the overall payoff has 4 possible values: $\{1, \alpha, (1 - \alpha), 0\}$ which one of them directly specify one of the 4 possible states realizations, that is, respectively. $\{(y^1 = 1, y^2 = 1), (y^1 = 1, y^2 = 0), (y^1 = 0, y^2 = 1), (y^1 = 0, y^2 = 0)\}$. However, there are important exceptions, when $\alpha = 1/2$ differentiate between cases 2 and 3, when $\alpha = 1$ we don't observe y^2 and when $\alpha = 0$ we don't observe y^1

Let y denote the overall payoff, we have the follow rule:

$$\begin{aligned} P(\theta = 1|y = 1) &= \frac{P(y = 1|\theta = 1)\mu}{P(y = 1)} \\ &= \frac{P(y^1 = 1, y^2 = 1|\theta = 1)\mu}{P(y^1, y^2 = 1)} \\ &= \frac{3/4 * 1/4\mu}{(3/4)(1/4)\mu + (1/4)(3/4)(1 - \mu)} \\ &= \frac{3/16\mu}{3/16} \\ &= \mu \end{aligned}$$

Similarly:

$$\begin{aligned} P(\theta = 1|y = \alpha) &= \frac{9\mu}{8\mu + 1} \\ P(\theta = 1|y = 1 - \alpha) &= \frac{\mu}{9 - 8\mu} \\ P(\theta = 1|y = 0) &= \mu \end{aligned} \quad (11)$$

But, if $\alpha = 1/2$ we have:

$$\begin{aligned} P(\theta = 1|y = 1/2) &= \frac{[P(y^1 = 1, y^0 = 0|\theta = 1) + P(y^1 = 0, y^0 = 1|\theta = 1)]\mu}{P(y^1 = 1, y^0 = 0) + P(y^1 = 1, y^0 = 0)} \\ &= \frac{10\mu}{(1 + 2\mu)^2 + (3 - 2\mu)^2} \end{aligned}$$

Finally, and most importantly, if $\alpha = 1$:

$$\begin{aligned} P(\theta = 1|y = 1) &= \frac{3\mu}{1 + 2\mu} \\ P(\theta = 1|y = 0) &= \frac{\mu}{3 - 2\mu} \end{aligned}$$

symmetrically, if $\alpha = 0$:

$$\begin{aligned} P(\theta = 1|y = 1) &= \frac{\mu}{3 - 2\mu} \\ P(\theta = 1|y = 0) &= \frac{3\mu}{1 + 2\mu} \end{aligned}$$

The agent choose alpha to maximize:

$$\begin{aligned}
& \max_{\alpha \in [0,1]} \alpha \mathbb{E}(y^1) + (1 - \alpha) \mathbb{E}(y^2) \\
& \max_{\alpha \in [0,1]} \alpha(1/2\mu + 1/4) + (1 - \alpha)(3/4 - 1/2\mu) \\
& \max_{\alpha \in [0,1]} \alpha(\mu - 1/2) + (3/4 - 1/2\mu)
\end{aligned} \tag{12}$$

So it is clear that as long as $\mu > 1/2$ the best choice is $\alpha = 1$, and as long as $\mu < 1/2$ the best choice is $\alpha = 0$, is $\mu = 1/2$, the agent is indifferent between all α .

With $\alpha = 1$ there is only two outcomes and updating rules. We can now use the same method as Q1, defining s_t as the difference between the number of positive outcomes and the number of negative outcomes. And $P_\infty(s)$ as the probability of playing α forever at state s .

As before, since $\mu \rightarrow 1$ as $s \rightarrow \infty$ we have that $P_\infty(\infty) = 1$. Therefore we can use exactly the same argument as Q1 to prove that $P_\infty(0) > 0$.