4.1-a. Objective function → $\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\max}\, \mathbb{E}_{\pi_\phi(y|x)}\left[r^*(x,y) - \beta \log \frac{\pi_\phi(y|x)}{\pi_{ref}(y|x)}\right]$

For changing from argmax to argmin we multiply the whole term by -1

So, $\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\min}\, \mathbb{E}_{\pi_\phi(y|x)}\left[\beta \log \frac{\pi_\phi(y|x)}{\pi_{ref}(y|x)} - r^*(x,y)\right]$

Now if we multiply β in denom & num.

So,

$$\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\min}\, \mathbb{E}_{\pi_\phi(y|x)}\left[\log \frac{\pi_\phi(y|x)}{\pi_{ref}(y|x)} - \frac{1}{\beta} r^*(x,y)\right]$$

On using the log property, we do exp & log on $\frac{1}{\beta} r^*(x,y)$

$$\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\min}\, \mathbb{E}_{\pi_\phi(y|x)}\left[\log \frac{\pi_\phi(y|x)}{\pi_{ref}(y|x)} - \log\left(\exp\left(\frac{1}{\beta} r^*(x,y)\right)\right)\right]$$

Using log property,

$$\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\min}\, \mathbb{E}_{\pi_\phi(y|x)}\left[\log \frac{\pi_\phi(y|x)}{\pi_{ref}(y|x)\exp\left(\frac{1}{\beta} r^*(x,y)\right)}\right]$$

We consider Z as normalization constant so divide Z by num & denom

So,

$$\pi_\phi^*(y|x) = \underset{\pi_\phi(y|x)}{\arg\min}\, \mathbb{E}_{\pi_\phi(y|x)}\left[\log \underbrace{\frac{\pi_\phi(y|x)/Z}{\left(\pi_{ref}(y|x)\exp\left(\frac{1}{\beta} r^*(x,y)\right)\right)/Z}}_{\text{This is KL divergence.}}\right]$$

As the $\log \frac{1}{Z}$ is constant. So, this is $+C$

Hence, when $\pi_\phi^*(y|x) = \pi_\phi(y|x)$ then KL divergence is minimized to 0.

i.e., both the distributions are equal.

So, we get

$$\boxed{\pi_\phi^*(y|x) = \frac{1}{Z}\pi_{ref}(y|x)\exp\left(\frac{1}{\beta} r^*(x,y)\right)}$$

Hence Proved.