Assignment 16

Question 3

Part 1

$$\nabla_\theta \log \pi(s,a;\theta) = \nabla_\theta \left(\phi(s,a)^T \cdot \theta - \log \sum_{b \in A} e^{\phi(s,b)^T \cdot \theta} \right)$$

$$= \phi(s,a) - \frac{\sum_{b \in A} \phi(s,b)\, e^{\phi(s,b)^T \cdot \theta}}{\sum_{b \in A} e^{\phi(s,b)^T \theta}}$$

$$= \phi(s,a) - \sum_{b \in A} \phi(s,b)\, \pi(s,b;\theta)$$

$$= \phi(s,a) - E_{\pi(s,b;\theta)}[\phi(s,b)]$$

Part 2

$$\nabla_w Q(s,a;w) = \nabla_\theta \log \pi(s,a;\theta)$$

We can easily set $Q(s,a;w) = w^T \nabla_\theta \log \pi(s,a;\theta)$

Part 3

$$E[Q(s,a;w)] = \sum_{a \in A} \pi(s,a;\theta)\, Q(s,a;w)$$

$$= \sum_{a \in A} \pi(s,a;\theta)\, w^T \left(\frac{1}{\pi(s,a;\theta)} \nabla_\theta \pi(s,a;\theta) \right)$$

$$= w^T \nabla_\theta \sum_{a \in A} \pi(s,a;\theta)$$

$$= w^T \nabla_\theta 1$$

$$= 0$$

Hence proved