Chapter 9

Exercise 9.1:

The general form of tabular updates is:
$v_\pi(S_t) = v_\pi(S_t) + \alpha[G - v_\pi(S_t)]$
Where G is the corresponding return.

The update for linear function approximation is:
$w_t = w_t + \alpha[G - w_t{}^T x(S_t)]x(S_t)$

Let w_t be a vector of N elements. Where N is the number of states.
$w_t = [v_\pi(0),\ v_\pi(1),\ v_\pi(2)....,v_\pi(N-1)]^T$
And $x_t(S_t) = [0, 0, 0, ...1, 0, 0, 0]^T$ such that there are S_t number of zeros before the one.

The update for linear approximation is now the same as the update for tabular methods. Only the value of S_t will change at time step t.

Exercise 9.2:

$x_i(s) = \prod_{j=1}^{k} s_j{}^{c_{i,j}}$ where $c_{i,j}$ can take any value from 0 to n.

For each $c_{i,j}$ we have $n+1$ choices

For $j = 1,\ s_1{}^{c_{i,1}}$ can have $n+1$ different values.

For $j = 2,\ s_2{}^{c_{i,2}}$ can have $n+1$ different values.

For $j = k,\ s_k^{c_{i,k}}$ can have $n + 1$ different values.

Therefore

$\prod_{j=1}^{k} s_j^{c_{i,j}}$ can have $\prod_{j=1}^{k}(n + 1)$ different values. Which is simply $(n + 1)^k$

Exercise 9.3:

$x(s) = (1,\ s_1,\ s_2,\ s_1 s_2,\ s_1^2,\ s_2^2,\ s_1 s_2^2,\ s_1^2 s_2,\ s_1^2 s_2^2)^T$

$k = 2$

$n = 2$ Since highest power is 2

$i = 1\ to\ 9$

$j = 1\ to\ 2$

$$
\begin{aligned}
c_{i,j} = {} & [0,\ 0] \\
& [1,\ 0] \\
& [0,\ 1] \\
& [1,\ 1] \\
& [2,\ 0] \\
& [0,\ 2] \\
& [1,\ 2] \\
& [2,\ 1] \\
& [2,\ 2]
\end{aligned}
$$

Exercise 9.4:

Since we want generalization across a dimension, any tiling which gives a larger area to the other dimension will be good. For example a rectangular tiling with the shorter side along the important dimension.

Exercise 9.5:

$\alpha = (\tau * E[x^T x])^{-1}$
$\tau = 10$ according to the exercise.