1.

a. There are three states (S’, S’’, and S’’’). The rewards for state S’, S’’, and S’’’ are -1, -2, and 10 respectively.

b. T(S’, a1, S’) = 0.25, T(S’, a1, S’’) = 0.75, T(S’’, a2, S’’) = 0.40, T(S’’, a1, S’’’) = 0.60

c.

Let γ = 1, δ = 0, ε = 0.01.

U(S’’) = R(S’’) + γ max ∑ P(s’ | s, a) U(s’)

U(S’’) = R(S’’) + γ max (0.4 U(S’’), 0.6 U(S’’’))

U(S’’) = R(S’’) + (1) max (0.4 U(S’’), 0.6 U(S’’’))

U(S’’) = R(S’’) + 0.6(U(S’’’))

U(S’’) = -2 + 0.6(10) = -2 + 6 = 4

U’(s) – U(s) = 4 – 0 = 4 > δ = 0, Therefore, δ = 4.

Stop if δ < ε => 4 < 0.01. Therefore, false.

U = U’. δ = 0.

U(S’’) = -2 + 0.6(10) = -2 + 6 = 4

U’(s) – U(s) = 4 – 4 = 0 > δ = 0. False.

Stop if δ < ε => 0 < 0.01. Therefore, true.

Therefore, U(S’’) = 4.

d. In this case you can, because the utility of U(S’’’) [which is the maximum for U(S’’)] will always be 10. As such, there is no need to used the value iterative algorithm as U’(S’’) will not change, the Bellman Equation will suffice.

e.

Let α = 0.1, γ = 1.

Episode 1:

Sample = R(S’’, a, S’’’) + γ V(S’’’) = 10 + (1)(10) = 20

V(S’’) = (1- α) V(S’’) + α(Sample) = (0.9)(-2) + (0.1)(20) = -1.8 + 2 = 0.2.

Episode 2:

Sample = R(S’’, a, S’’) + γ V(S’’) = -2 + (1)(0.2) = -1.8

V(S’’) = (1- α) V(S’’) + α(Sample) = (0.9)(0.2) + (0.1)(-1.8) = 0.18 – 0.18 = 0.

Sample = R(S’’, a, S’’’) + γ V(S’’’) = 10 + (1)(10) = 20

V(S’’) = (1- α) V(S’’) + α(Sample) = (0.9)(0) + (0.1)(20) = 0 + 2 = 2.

Episode 3:

Sample = R(S’’, a, S’’) + γ V(S’’) = -2 + (1)(2) = 0

V(S’’) = (1- α) V(S’’) + α(Sample) = (0.9)(2) + (0.1)(0) = 1.8 + 0 = 1.8.

Sample = R(S’’, a, S’’’) + γ V(S’’’) = 10 + (1)(10) = 20

V(S’’) = (1- α) V(S’’) + α(Sample) = (0.9)(1.8) + (0.1)(20) = 1.62 + 2 = 3.62.

Therefore, V(S’’) = 3.62.

2. Prove if P(X | Y, Z) = P(Y | X, Z), then P(X | Z) = P(Y | Z)

P(X | Y, Z) = P(Y | X, Z)

P(X, Y, Z) / P(Y, Z) = P(Y, X , Z) / P(X, Z)

P(X, Y, Z) / P(Y, Z) = P(X, Y , Z) / P(X, Z)

1 / P(Y, Z) = 1 / P(X, Z)

P(X, Z) = P(Y, Z)

P(X | Z) \* P(Z) = P(Y | Z) \* P(Z)

P(X | Z) = P(Y | Z).

Hence, proved.

3.

a. P(A, C, D, B!, E!) = P(A) P(C | A) P(D | B, C) P (!B | !A) P(!E | !C).

b. Markov blanket of a variable is a set containing the variable’s parents, children, and co-parents. As such, the Markov blanket separates the variable from all other variables, hence, making the variable independent of all other variables. This is because any path from the variable to outside its Markov blanket will have to connect to another variable before leaving the blanket. This path would also either be serial or diverging. As such, we can say that a variable is independent of all other variables except its Markov blanket.

c.

P(D | E, !B)

= P(D, E, !B) / P(E, !B)

= α P(D, E, !B)

= α ∑A ∑C P(D, A, C, E, !B)

= α ∑A ∑C P(D | B, C) P(A) P(C | A) P(E | C) P(!B | !A)