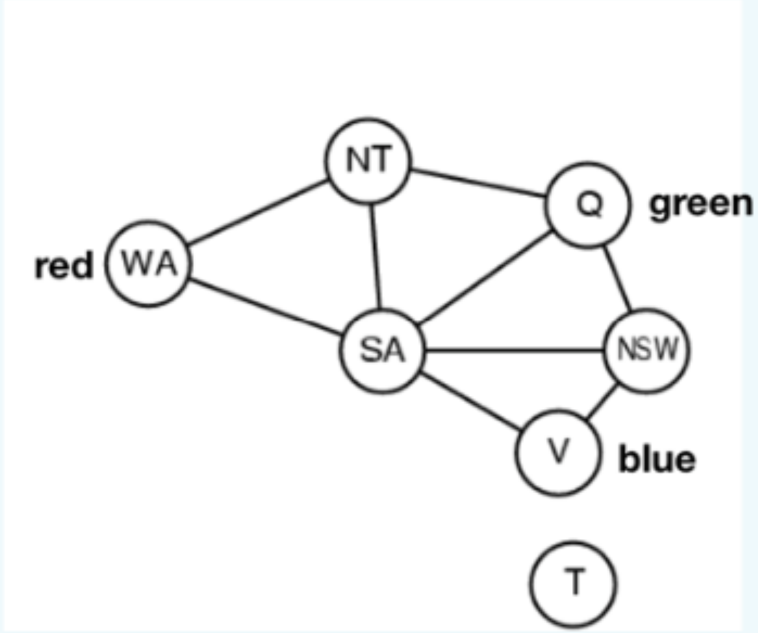
COMP3411/9814

23T1

QUIZ 2

Question 1

Consider the constraint satisfaction problem (CSP) of colouring the Australian map with three colours and suppose Westerr Australia has been assigned red (WA = {red}), Queensland green (Q = {green}) and Victoria blue (V = {blue}), as shown below

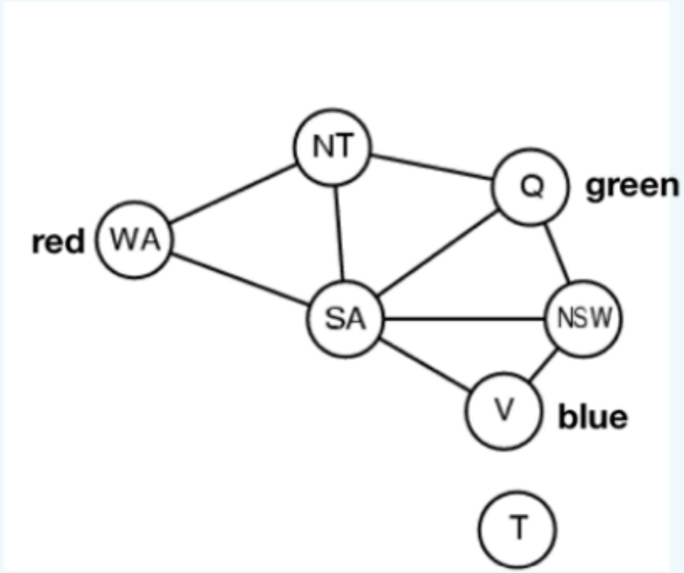


What are the remaining values for the other variables after applying **forward checking**?

- [ ] NT = {blue}, NSW = {blue}, SA = {green}, TAS = {red, blue, green}
- [ ] NT = {blue}, NSW = {red}, SA = {blue}, TAS = {red, blue, green}
- [x] NT = {blue}, NSW = {red}, SA = {}, TAS = {red, blue, green}
- [ ] NT = {blue}, NSW = {}, SA = {red}, TAS = {red, blue, green}

Question 2

Consider the constraint satisfaction problem (CSP) of colouring the Australian map with three colours and suppose Western Australia has been assigned red (WA = {red}), Queensland green (Q = {green}) and Victoria blue (V = {blue}), as shown below.

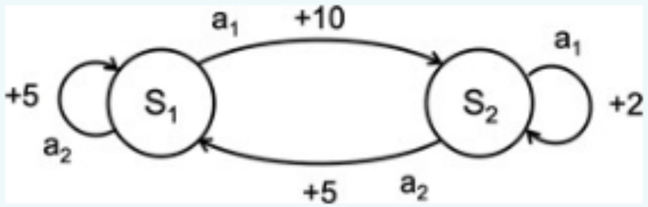


What are the remaining values for the other variables after applying **arc consistency checking**?

- ○ NT = {blue}, NSW = {blue}, SA = {green}, TAS = {red, blue, green}
- ○ NT = {blue}, NSW= {}, SA = {red}, TAS = {red, blue, green}
- ● NT = {blue}, NSW = {red}, SA = {}, TAS = {red, blue, green}
- ○ NT = {blue}, NSW= {red}, SA = {blue}, TAS = {red, blue, green}

Question 3

Consider the following reinforcement learning problem.

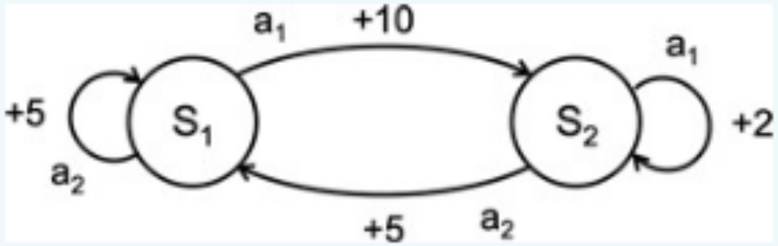


Which relationship holds of the optimal value function V^* under infinite discounted rewards with $\gamma = 0.8$?

- ○ $V^*(S_1) = 5 + V^*(S_1)$
- ○ $V^*(S_1) = 5 + 0.8 * V^*(S_1)$
- ○ $V^*(S_1) = 10 + 0.8 * (2 + 0.8 * V^*(S_2))$
- ● $V^*(S_1) = 10 + 0.8 * V^*(S_2)$

Question 4

What is the optimal policy π^* for the reinforcement learning problem below ?



- ( ) $\pi^*(S_1) = a_2, \pi^*(S_2) = a_1$
- (•) $\pi^*(S_1) = a_1, \pi^*(S_2) = a_2$
- ( ) $\pi^*(S_1) = a_1, \pi^*(S_2) = a_1$
- ( ) $\pi^*(S_1) = a_2, \pi^*(S_2) = a_2$

Question 5

Consider the discounted return equation (1) with a discount factor $\gamma = 0.5$ and a reward sequence of 10, 20, 1, 10, 100. The returns G_0 and G_1 are equal to:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \ldots = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{1}$$

- (selected) $G_0 = 27.75$, $G_1 = 35.5$
- $G_0 = 20.25$, $G_1 = 27.75$
- $G_0 = 0$, $G_1 = 10$
- $G_0 = 10$, $G_1 = 20$