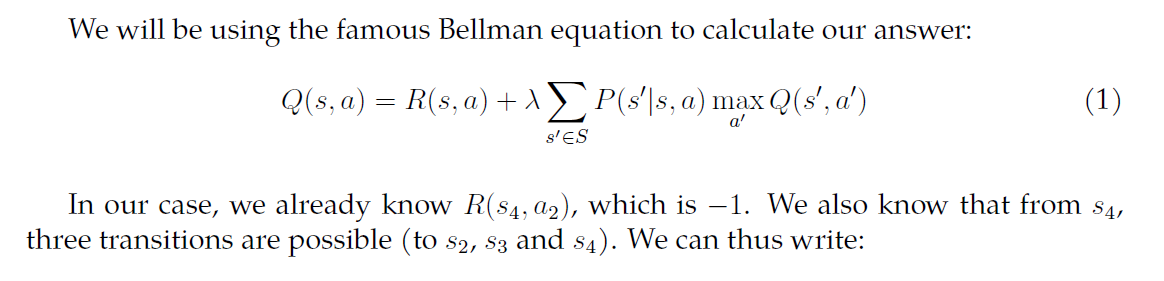
# Solutions to the exercises on reinforcement learning and dialogue system evaluation

# (IN4080 course)

## Question 1

1.1



## 

1.2



## Question 2

We need to compute two Q-values:

So, for the first iteration, starting with Q-values initialized to 0, we get:

For the second iteration, we get:

We continue:

So the correct answer at the end of iteration fine would be 6.28 for Q(HumanNotEngaged, SayHi) and 5.97 for Q(HumanNotEngaged, SayHiWithGestures).