MDL Assignment 3 Part 1

Team Name: Trees.Love

Team Number: 5

Team Members:

Pulak Malhotra, 2019101050

Ashwin Rao, 2019101049

Belief Update

After reaching s', the agent observes $o \in \Omega$ with probability $O(o \mid s', a)$. Let b be a probability distribution over the state space S. $b(s)$ denotes the probability that the environment is in state s. Given $b(s)$, then after taking action a and observing o,

$$b'(s') = \eta O(o \mid s', a) \sum_{s \in S} T(s' \mid s, a) b(s)$$

where $\eta = 1/\Pr(o \mid b, a)$ is a normalizing constant

Calculations:

Move is R, Observation is G

Calculating b_prime[0]

- T(0→0,R) = 0.01
  - On multiplying by b[0]=0.3333333 gives 0.0033333

- T(1→0,R) = 0.01
    - On multiplying by b[1]=0 gives 0.0
- T(2→0,R) = 0
- T(3→0,R) = 0
- T(4→0,R) = 0
- T(5→0,R) = 0
  sum of transitions 0.003333333333333336
  P(o=G|a=R,s'=0) = 0.15000000000000002
  b_prime[0]= 0.0033333 * 0.15 = 0.0005

Calculating b_prime[1]

- T(0→1,R) = 0.99
    - On multiplying by b[0]=0.3333333 gives 0.33
- T(1→1,R) = 0
- T(2→1,R) = 0.01
    - On multiplying by b[2]=0.3333333 gives 0.0033333
- T(3→1,R) = 0
- T(4→1,R) = 0
- T(5→1,R) = 0
  sum of transitions 0.3333333333333333
  P(o=G|a=R,s'=1) = 0.9
  b_prime[1]= 0.3333333 * 0.9 = 0.3

Calculating b_prime[2]

- T(0→2,R) = 0
- T(1→2,R) = 0.99
    - On multiplying by b[1]=0 gives 0.0
- T(2→2,R) = 0
- T(3→2,R) = 0.01

    - On multiplying by b[3]=0 gives 0.0
- T(4→2,R) = 0
- T(5→2,R) = 0
  sum of transitions 0.0
  P(o=G|a=R,s'=2) = 0.15000000000000002
  b_prime[2]= 0.0 * 0.15 = 0.0

Calculating b_prime[3]

- T(0→3,R) = 0
- T(1→3,R) = 0
- T(2→3,R) = 0.99
    - On multiplying by b[2]=0.3333333 gives 0.33
- T(3→3,R) = 0
- T(4→3,R) = 0.01
    - On multiplying by b[4]=0 gives 0.0
- T(5→3,R) = 0
  sum of transitions 0.32999999999999996
  P(o=G|a=R,s'=3) = 0.9
  b_prime[3]= 0.33 * 0.9 = 0.297

Calculating b_prime[4]

- T(0→4,R) = 0
- T(1→4,R) = 0
- T(2→4,R) = 0
- T(3→4,R) = 0.99
    - On multiplying by b[3]=0 gives 0.0
- T(4→4,R) = 0
- T(5→4,R) = 0.01
    - On multiplying by b[5]=0.3333333 gives 0.0033333
      sum of transitions 0.003333333333333336

P(o=G|a=R,s'=4) = 0.9
b_prime[4]= 0.0033333 * 0.9 = 0.003

Calculating b_prime[5]

- T(0→5,R) = 0
- T(1→5,R) = 0
- T(2→5,R) = 0
- T(3→5,R) = 0
- T(4→5,R) = 0.99
  - On multiplying by b[4]=0 gives 0.0
- T(5→5,R) = 0.99
  - On multiplying by b[5]=0.3333333 gives 0.33
    sum of transitions 0.32999999999999996
    P(o=G|a=R,s'=5) = 0.15000000000000002
    b_prime[5]= 0.33 * 0.15 = 0.0495

Not normalized numerator
[0.0005000000000000004, 0.3, 0.0, 0.297, 0.0030000000000000027, 0.0495]
Denominator 0.6499999999999999
Updated Beliefs
[0.00076923076923077, 0.46153846153846156, 0.0, 0.456923076923077, 0.00461538461538462, 0.07615384615384617]

Move is L, Observation is R

Calculating b_prime[0]

- T(0→0,L) = 0.99
  - On multiplying by b[0]=0.0007692 gives 0.0007615
- T(1→0,L) = 0.99
  - On multiplying by b[1]=0.4615385 gives 0.4569231
- T(2→0,L) = 0
- T(3→0,L) = 0

- T(4→0,L) = 0
- T(5→0,L) = 0
  sum of transitions 0.4576846153846154
  P(o=R|a=L,s'=0) = 0.85
  b_prime[0]= 0.4576846 * 0.85 = 0.3890319

Calculating b_prime[1]

- T(0→1,L) = 0.01
  - On multiplying by b[0]=0.0007692 gives 7.7e-06
- T(1→1,L) = 0
- T(2→1,L) = 0.99
  - On multiplying by b[2]=0.0 gives 0.0
- T(3→1,L) = 0
- T(4→1,L) = 0
- T(5→1,L) = 0
  sum of transitions 7.692307692307707e-06
  P(o=R|a=L,s'=1) = 0.09999999999999998
  b_prime[1]= 7.7e-06 * 0.1 = 8e-07

Calculating b_prime[2]

- T(0→2,L) = 0
- T(1→2,L) = 0.01
  - On multiplying by b[1]=0.4615385 gives 0.0046154
- T(2→2,L) = 0
- T(3→2,L) = 0.99
  - On multiplying by b[3]=0.4569231 gives 0.4523538
- T(4→2,L) = 0
- T(5→2,L) = 0
  sum of transitions 0.456969230769230 85

P(o=R|a=L,s'=2) = 0.85
b_prime[2]= 0.4569692 * 0.85 = 0.3884238

Calculating b_prime[3]

- T(0→3,L) = 0
- T(1→3,L) = 0
- T(2→3,L) = 0.01
  - On multiplying by b[2]=0.0 gives 0.0
- T(3→3,L) = 0
- T(4→3,L) = 0.99
  - On multiplying by b[4]=0.0046154 gives 0.0045692
- T(5→3,L) = 0
  sum of transitions 0.004569230769230774
  P(o=R|a=L,s'=3) = 0.09999999999999998
  b_prime[3]= 0.0045692 * 0.1 = 0.0004569

Calculating b_prime[4]

- T(0→4,L) = 0
- T(1→4,L) = 0
- T(2→4,L) = 0
- T(3→4,L) = 0.01
  - On multiplying by b[3]=0.4569231 gives 0.0045692
- T(4→4,L) = 0
- T(5→4,L) = 0.99
  - On multiplying by b[5]=0.0761538 gives 0.0753923
    sum of transitions 0.07996153846153847
    P(o=R|a=L,s'=4) = 0.09999999999999998
    b_prime[4]= 0.0799615 * 0.1 = 0.0079962

Calculating b_prime[5]

- T(0→5,L) = 0

- T(1→5,L) = 0
- T(2→5,L) = 0
- T(3→5,L) = 0
- T(4→5,L) = 0.01
  - On multiplying by b[4]=0.0046154 gives 4.62e-05
- T(5→5,L) = 0.01
  - On multiplying by b[5]=0.0761538 gives 0.0007615
    sum of transitions 0.0008076923076923086
    P(o=R|a=L,s'=5) = 0.85
    b_prime[5]= 0.0008077 * 0.85 = 0.0006865

Not normalized numerator
[0.3890319230769231, 7.692307692307705e-07, 0.38842384615384623, 0.0004569230769230773, 0.007996153846153846, 0.0006865384615384623]
Denominator 0.786596153846154
Updated Beliefs
[0.494576436936166, 9.779233797032017e-07, 0.49380338850451067, 0.0005808864875437013, 0.010165513532014763, 0.0008727966616385107]

Move is L, Observation is G

Calculating b_prime[0]

- T(0→0,L) = 0.99
  - On multiplying by b[0]=0.4945764 gives 0.4896307
- T(1→0,L) = 0.99
  - On multiplying by b[1]=1e-06 gives 1e-06
- T(2→0,L) = 0
- T(3→0,L) = 0
- T(4→0,L) = 0
- T(5→0,L) = 0
  sum of transitions 0.4896316407109502

P(o=G|a=L,s'=0) = 0.15000000000000002
b_prime[0]= 0.4896316 * 0.15 = 0.0734447

Calculating b_prime[1]

- T(0→1,L) = 0.01
  - On multiplying by b[0]=0.4945764 gives 0.0049458
- T(1→1,L) = 0
- T(2→1,L) = 0.99
  - On multiplying by b[2]=0.4938034 gives 0.4888654
- T(3→1,L) = 0
- T(4→1,L) = 0
- T(5→1,L) = 0
  sum of transitions 0.49381111898882724
  P(o=G|a=L,s'=1) = 0.9
  b_prime[1]= 0.4938111 * 0.9 = 0.44443

Calculating b_prime[2]

- T(0→2,L) = 0
- T(1→2,L) = 0.01
  - On multiplying by b[1]=1e-06 gives 0.0
- T(2→2,L) = 0
- T(3→2,L) = 0.99
  - On multiplying by b[3]=0.0005809 gives 0.0005751
- T(4→2,L) = 0
- T(5→2,L) = 0
  sum of transitions 0.0005750874019020613
  P(o=G|a=L,s'=2) = 0.15000000000000002
  b_prime[2]= 0.0005751 * 0.15 = 8.63e-05

Calculating b_prime[3]

- T(0→3,L) = 0

- T(1→3,L) = 0
- T(2→3,L) = 0.01
  - On multiplying by b[2]=0.4938034 gives 0.004938
- T(3→3,L) = 0
- T(4→3,L) = 0.99
  - On multiplying by b[4]=0.0101655 gives 0.0100639
- T(5→3,L) = 0
  sum of transitions 0.015001892281739725
  P(o=G|a=L,s'=3) = 0.9
  b_prime[3]= 0.0150019 * 0.9 = 0.0135017

Calculating b_prime[4]

- T(0→4,L) = 0
- T(1→4,L) = 0
- T(2→4,L) = 0
- T(3→4,L) = 0.01
  - On multiplying by b[3]=0.0005809 gives 5.8e-06
- T(4→4,L) = 0
- T(5→4,L) = 0.99
  - On multiplying by b[5]=0.0008728 gives 0.0008641
    sum of transitions 0.000869877515096693
    P(o=G|a=L,s'=4) = 0.9
    b_prime[4]= 0.0008699 * 0.9 = 0.0007829

Calculating b_prime[5]

- T(0→5,L) = 0
- T(1→5,L) = 0
- T(2→5,L) = 0
- T(3→5,L) = 0

- T(4→5,L) = 0.01
  - On multiplying by b[4]=0.0101655 gives 0.0001017
- T(5→5,L) = 0.01
  - On multiplying by b[5]=0.0008728 gives 8.7e-06
    sum of transitions 0.0001103831014839988
    P(o=G|a=L,s'=5) = 0.15000000000000002
    b_prime[5]= 0.0001104 * 0.15 = 1.66e-05

Not normalized numerator
[0.07344474610664255, 0.4444300070899445, 8.62631102853092e-05, 0.013501703053565753, 0.0007828897635870237, 1.6557465222599824e-05]
Denominator 0.5322621665892476
Updated Beliefs
[0.13798603529775325, 0.8349832751365097, 0.00016206883693816877, 0.025366640541230053, 0.00147087246235028 4, 3.110772521875934e-05]