CS4100 Problem Set 2

Silas Nevstad

July 2024

1. In this model we can see that it is both irreducible (any two states are able to reach each other) and aperiodic (transitions do not follow a fixed cycle; S1 has a self-transition and S2/S3 can reach themselves in 2 or 3 steps). These two properties guarantee that there exists a unique stationary distribution. Therefore, this stationary distribution does not depend on the initial state because, over time, the system will converge to this distribution.

2. $\pi = [0.3864, 0.4091, 0.2045]$

3. $P(S1 \to S1 \to S2 \to S3) = \pi(S1) \times P(S1 \to S1) \times P(S1 \to S2) \times P(S2 \to S3)$

   $P(S1 \to S1 \to S2 \to S3) = 0.3864 \times 0.1 \times 0.9 \times 0.5$

   $P(S1 \to S1 \to S2 \to S3) = 0.0174$

4. $P(S2 \to S2, intwosteps) = P(S2 \to S1 \to S2) + P(S2 \to S2 \to S2) + P(S2 \to S3 \to S2)$

   $P(S2 \to S2, intwosteps) = (P(S2 \to S1) \times P(S1 \to S2)) + (P(S2 \to S2) \times P(S2 \to S2)) + (P(S2 \to S3) \times P(S3 \to S2))$

   $P(S2 \to S2, intwosteps) = (0.5 \times 0.9) + (0 \times 0) + (0.5 \times 0.3)$

   $P(S2 \to S2, intwosteps) = 0.45 + 0 + 0.15$

   $P(S2 \to S2, intwosteps) = 0.6$

5. $p_0 = [0.1, 0.6, 0.3]$

   $$P = \begin{bmatrix} 0.1 & 0.9 & 0.0 \\ 0.5 & 0.0 & 0.5 \\ 0.7 & 0.3 & 0.0 \end{bmatrix}$$

   $p_2 = p_0 \times P \times P$

   $p_2 = [0.52, 0.18, 0.3] \times P$

   $p_2 = [0.352, 0.558, 0.09]$

   So the probability of ending up in S2 after exactly two transitions is 0.558.

6.

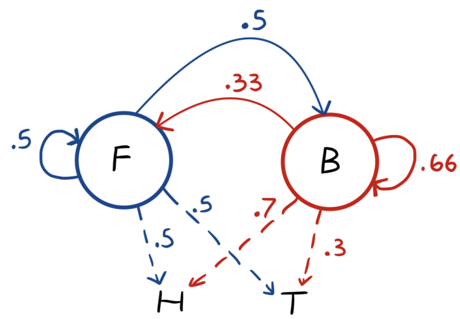

Figure 1: HMM Diagram for coin flipping game

7. $P = \begin{bmatrix} 0.5 & 0.5 \\ 0.33 & 0.66 \end{bmatrix}$

   $\pi = [0.3976, 0.6024]$

   $\delta_1(F) = \pi(F) \times P(H \mid F) = 0.3976 \times 0.5 = 0.1988$

   $\delta_1(B) = \pi(B) \times P(H \mid B) = 0.6024 \times 0.7 = 0.4217$

   $\delta_2(F) = max(0.1988 \times 0.5, 0.4217 \times 1/3) \times 0.5 = max(0.0994, 0.1406) \times 0.5 = 0.0703$

   $\delta_2(B) = max(0.1988 \times 0.5, 0.4217 \times 2/3) \times 0.7 = max(0.0994, 0.2811) \times 0.7 = 0.1968$

   $\delta_3(F) = max(0.0703 \times 0.5, 0.1968 \times 1/3) \times 0.5 = max(0.03515, 0.0656) \times 0.5 = 0.0328$

   $\delta_3(B) = max(0.0703 \times 0.5, 0.1968 \times 2/3) \times 0.3 = max(0.03515, 0.1312) \times 0.3 = 0.0394$

   Therefore, the most likely hidden sequence is $B \rightarrow B \rightarrow B$

8. - S1 has 2 actions
   - S2 has 1 action
   - S3 has 1 action
   - S4 has no actions

   2 * 1 * 1 = 2 unique policies possible

   $\pi_1 = S1 \rightarrow A1, S2 \rightarrow A3, S3 \rightarrow A4$

   $\pi_2 = S1 \rightarrow A2, S2 \rightarrow A3, S3 \rightarrow A4$

9. $Path_1 = S1 \rightarrow S3 \rightarrow S4$ via A2, A4
   $Path_2 = S1 \rightarrow S2 \rightarrow S3 \rightarrow S4$ via A1, A3, A4
   $Path_3 = S1 \rightarrow S2 \rightarrow S3 \rightarrow S4$ via A2, A3, A4

   $P(S1 \rightarrow S2, A1) = 0.6$
   $P(S1 \rightarrow S2, A2) = 0.8$
   $P(S1 \rightarrow S3, A2) = 0.2$
   $P(S2 \rightarrow S3) = 0.7$
   $P(S3 \rightarrow S4) = 0.7$

   $P(Path_1) = (0.5 \times 0.2) \times (1 \times 0.7) = 0.07$
   $P(Path_2) = (0.5 \times 0.6) \times (1 \times 0.7) \times (1 \times 0.7) = 0.147$
   $P(Path_3) = (0.5 \times 0.8) \times (1 \times 0.7) \times (1 \times 0.7) = 0.196$

   $P(S1 \rightarrow S4) = P(Path_1) + P(Path_2) + P(Path_3)$
   $P(S1 \rightarrow S4) = 0.07 + 0.147 + 0.196$
   $P(S1 \rightarrow S4) = 0.413$

10. $V_{opt}(S1, A1) = 0.6[10 + \gamma V_{opt}(S2)] + 0.4[0 + \gamma V_{opt}(S1)]$
    $V_{opt}(S1, A2) = 0.8[10 + \gamma V_{opt}(S2)] + 0.2[15 + \gamma V_{opt}(S3)]$
    $V_{opt}(S1) = max(0.6[10+\gamma V_{opt}(S2)]+0.4[\gamma V_{opt}(S1)], 0.8[10+\gamma V_{opt}(S2)]+ 0.2[15 + \gamma V_{opt}(S3)])$
    $V_{opt}(S1) = max(6 + 0.6\gamma V_{opt}(S2) + 0.4\gamma V_{opt}(S1), 11 + 0.8\gamma V_{opt}(S2) + 0.2\gamma V_{opt}(S3))$

11. $\eta = 1/(1 + 0) = 1$
    $Q(S1, A1) = (1 - 1)0 + 1(10 + 1 * max(0)) = 10$

    $\eta = 1/(1 + 0) = 1$
    $Q(S2, A3) = (1 - 1)0 + 1(0 + 1 * max(0)) = 0$

    $\eta = 1/(1 + 0) = 1$
    $Q(S3, A4) = (1 - 1)0 + 1(100 + 1 * 0) = 100$

    $\eta = 1/(1 + 0) = 1$
    $Q(S1, A2) = (1 - 1)0 + 1(15 + 1 * max(Q(S3, A4))) = 15 + 100 = 115$

    $Q_{opt}(S1, A1) = 10$
    $Q_{opt}(S1, A2) = 115$

$Q_{opt}(S2, A3) = 0$
$Q_{opt}(S3, A4) = 100$

$\pi(S1) = A2$
$\pi(S2) = A3$
$\pi(S3) = A4$

12. I have read and understood the academic integrity policy as outlined in the course syl- labus for CS4100. By pasting this acknowledgement in my submission, I declare that all work presented here is my own, and any conceptual discussions I may have had with classmates have been fully disclosed. I declare that generative AI was not used to answer any questions in this assignment. Any use of generative AI to improve writing clarity alone is accompanied by an appendix with my original, unedited answers.