35

Mixed Chains

In this chapter we learn how to analyze Markov chains that consists of transient and absorbing states. Later we will see that this analysis extends easily to chains with (non-absorbing) ergodic states. Unlike in the previous chapters, I will use an example to first teach the material. First I am going to outline the steps of our primary procedure.

We are given some Markov chain, A, whose states consists of transient and absorbing states. It is convenient in what follows that we renumber the vertices if necessary so that the absorbing states precede the transient states. We then partition the matrix as follows: $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{M} & \boldsymbol{T} \end{pmatrix}$. All four sub-matrices are rectangular, with I and T necessarily square. The square sub-matrix I corresponds to the transitions of absorbing states to absorbing states, and is thus an identity matrix. The square sub-matrix T corresponds to the transitions of the transient states to transient states. The sub-matrix O, represents the transitions of absorbing states to transient states and is thus all zeros. The sub-matrix M (for *mixed*) gives the transitions of transient states to absorbing states.

We will be interested in the matrix (I − T) where I is an identity matrix of the same dimension as T and is not the identity matrix I given earlier (which is likely to be dimensioned differently). We will find the inverse of the matrix $TT = (I - T)^{-1}$ which is denoted TT to indicate that it relates transient states to transient states. Multiplying TT on the left of M, we get the matrix $TA = TT \cdot M$, where the notation TA indicates a relation from transient states to absorbing states. The matrices TT and TA are the objects of our labors.

Penney-Ante

The game Penney-ante[1] is played by two players Alicia and Bernard. They flip a fair coin until they get a sequence of throws of the form heads-heads-tails in which case Alicia wins, or they get heads-tails-tails and Bernard wins. If we throw the coin three times, each has the same probability of winning (1/8). If neither wins we could follow with another three throws and continue in that manner until one of them wins. If the game is played that way, it is a fair game: each has a 50% probability of winning. However, Penney-ante is not played that way. If after the first three throws there is no winner, we do a fourth throw and we examine the last three throws. If there is still no winner, we do a fifth throw, and we continue in that manner until there is a win. To most people it still seems intuitively to be a fair game.

[1]Walter Penney. "Problem 95: Penney-Ante." *Journal of Recreational Mathematics*. **7** (1974), p. 321.

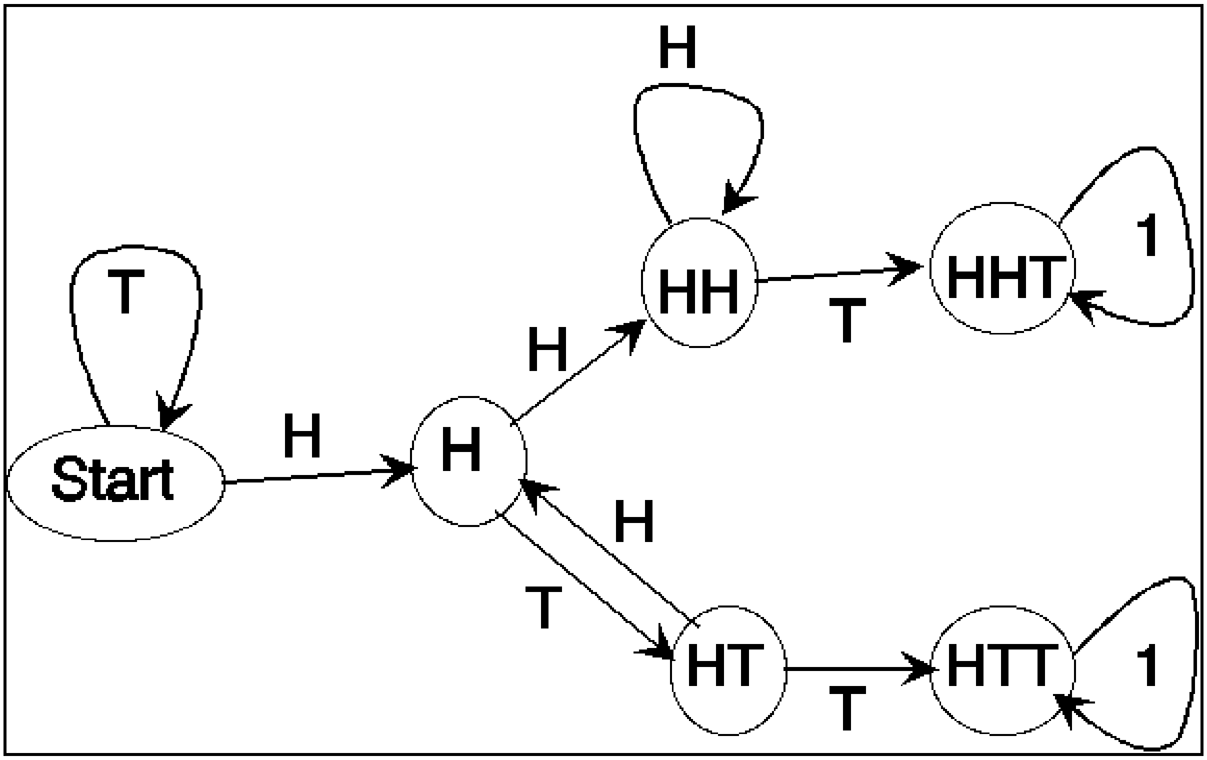

Figure 1 The game of Penney-Ante

Penney-ante is a game ideally suited to be represented by a Markov chain. The representation is given in **Figure 1**. The states HHT and HTT are absorbing states because they represent the end of the game. In the start state, throwing tails does not initiate any winning sequence, so for all practical purposes the game can be considered in the start state until heads is thrown. If the game is in state HT and heads is thrown, that heads becomes the possible first heads of a winning sequence. Clearly, once the game is in the state HH, only the sequence HHT can win. This is the key to the insight Alicia (who is betting on HHT) may have a significant advantage.

	HHT	HTT	Start	H	HT	HH
HHT	1	0	0	0	0	0
HTT	0	1	0	0	0	0
Start	0	0	.5	.5	0	0
H	0	0	0	0	.5	.5
HT	0	.5	0	.5	0	0
HH	.5	0	0	0	0	.5

The Transition Matrix for Penney-Ante

Following the course described above, we build the transition matrix by placing the absorbing states first followed by the transient states. The matrix is partitioned with M being a 4-by-2 sub-matrix and T being a 4-by-4 sub-matrix. We need the matrix $I-T$ and then its inverse, $(I-T)^{-1}$. The matrix $(I-T)^{-1}$, which we denote by TT has an interesting interpretation.

.5	-.5	0	0
0	1	-.5	-.5
0	-.5	1	0
0	0	0	.5

	Start	H	HT	HH
Start	2	1.333	.667	1.333
H	0	1.333	.667	1.333
HT	0	.667	1.333	.667
HH	0	0	0	2

The Matrix I-T and the Matrix TT = $(I-T)^{-1}$

The matrix TT is a matrix whose rows and columns correspond to transient states. The number in the box {m, n} (where m is the row and n is the column) represents the expected number of times the chain will visit state n having started in state m. For example, in the above case, once in state HH we can expect to visit it twice. However, that includes the visit of starting there. Hence the main diagonal of the TT matrix will contain entries greater or equal to 1. So in real terms, starting in state HH we can expect to return there once (on average). Starting in state H we can expect to visit state HT 4/3 times before being absorbed.

Lastly, we multiply the matrix TT times the matrix M (from the left).. This gives us the matrix TA. This is a matrix whose rows correspond to transient states and whose columns correspond to absorbing states. According to this matrix, on leaving the node we labeled Start, the probability of being eventually absorbed in HHT is .667 (or 2/3). That is, Alicia has a two-thirds probability of winning this game! Similarly, the absorption probabilities are the same at H as at Start. This is because on leaving Start the chain always goes to H. Also, the matrix verifies the earlier observation, that on reaching the node HH, the only winning (that is, absorbing) node possible is HHT. On the other hand, having reached HT, there is still a one-third probability of reaching HHT. Note, that in the TA matrix, the rows must always add up to 1.

	HHT	HTT
Start	.667	.333
H	.667	.333
HT	.333	.667
HH	1	0

The Matrix TA = TT·M

Compleat Craps

Craps is a dice game that is best analyzed as a Markov chain.[1] The rules of craps are fairly simple. Craps can be played by any number of people. The person throwing the dice is the *player*. The player throws an ordinary pair of dice. If the first throw yields a 7 or an 11, he wins his bet. If that first throw yields a 2, 3, or 12, he loses the bet. Any other throw, 4, 5, 6, 8, 9, or 10, establishes the *point*. For example, if the player's first roll is a 4, then his point is 4.

[1]Craps is a wholesome game that brings enjoyment to the whole family. It is best enjoyed when played for real money.

After that he keeps throwing the dice until he gets a 4 or a 7. Any other throw (such as 2, 3, 4, [1]5) is irrelevant. If he throws the point before 7 he wins the bet. If he throws 7 first, he *craps out*. It is not hard to show that the player's probability of winning the bet is .492929.... The Markov chain for craps is given in **Figure 2**.

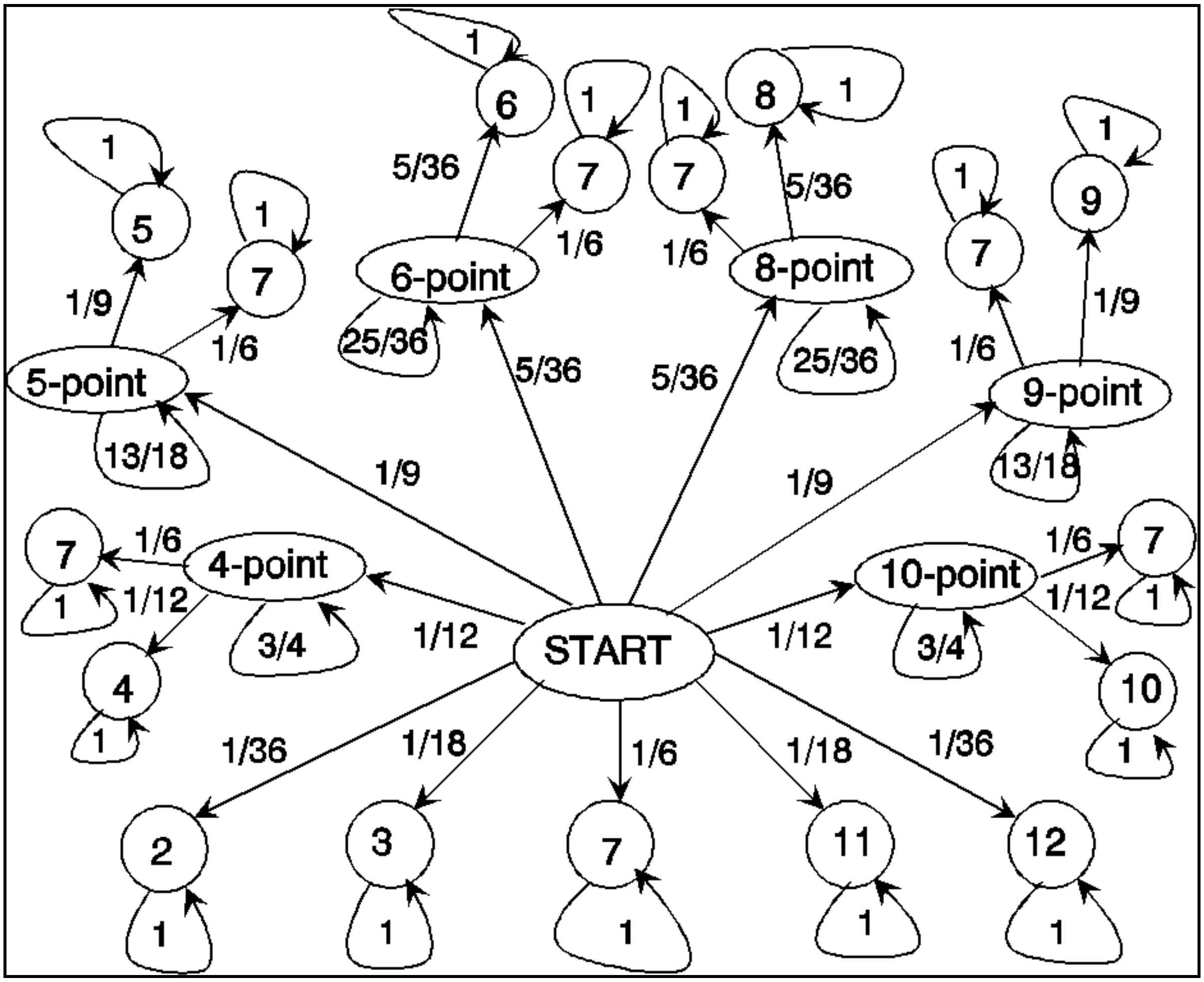

Figure 2 The Markov Chain for Craps

[1]Note that the last topic of Section 31, The Probability That A Occurs Before B, applies here. Using it we can look at the Markov chain for craps and deduce various probabilities. For example if the point is 4, the probability of throwing a 7 (and crapping out) is twice the probability of throwing a 4. Hence the probability of crapping out (with a 4 point) is ⅔.

Again, if the first throw is a 4, then that is the point. After that the player keeps throwing until he gets a 4 or a 7. He has, in that instance, a probability of 3/4 of throwing a value other than 4 or 7. He has a 1/12 probability of throwing a 4 and a 1/6 probability of throwing a 7. The TT matrix for craps is given below.

	Start	4-p	5-p	6-p	8-p	9-p	10-p
Start	1.000	0.333	0.400	0.455	0.455	0.400	0.333
4-p	0.000	4.000	0.000	0.000	0.000	0.000	0.000
5-p	0.000	0.000	3.600	0.000	0.000	0.000	0.000
6-p	0.000	0.000	0.000	3.273	0.000	0.000	0.000
8-p	0.000	0.000	0.000	0.000	3.273	0.000	0.000
9-p	0.000	0.000	0.000	0.000	0.000	3.600	0.000
10-p	0.000	0.000	0.000	0.000	0.000	0.000	4.000

The TT Matrix for Craps

We will use this matrix to calculate the expected number of throws in a game of craps. The probability that the game of craps ends in one throw, is the probability that the player throws a 2, 3, 7, 11, or 12. That probability is 1/3. The other probability is that the player establishes a point. We know that the probability that the player throws a 4 on his first throw is 1/12. According to the matrix, the expected number of trips to transient states before being absorbed is 4, which includes the initial visit to that state. The expected number of throws for craps in general is the summation of the probability of each outcome of the first throw times the expected number of throws in each case:

$$P(2)1 + P(3)1 + P(4)4 + P(5)3.6 + P(6)3.273 + P(7)1 + P(8)3.273 + P(9)3.6 + P(10)4 + P(11)1 + P(12)1 =$$
$$\frac{1}{36}1 + \frac{2}{36}1 + \frac{3}{36}4 + \frac{4}{36}3.6 + \frac{5}{36}3.273 + \frac{6}{36}1 + \frac{5}{36}3.273 + \frac{4}{36}3.6 + \frac{3}{36}4 + \frac{2}{36}1 + \frac{1}{36}1 = 2.7091\overline{6}$$

Hence, in a game of craps the expected number of throws is roughly 2.7.

To complete our analysis of craps, we examine the matrix TA = TT·M.

	2	3	7	11	12
Start	.028	.056	.167	.056	.028
4-p	.000	.000	.000	.000	.000
5-p	.000	.000	.000	.000	.000
6-p	.000	.000	.000	.000	.000
8-p	.000	.000	.000	.000	.000
9-p	.000	.000	.000	.000	.000
10-p	.000	.000	.000	.000	.000

	4-4	4-7	5-5	5-7	6-6	6-7	8-8	8-7	9-9	9-7	10-10	10-7
Start	.028	.056	.044	.067	.063	.076	.063	.076	.044	.067	.028	.056
4-p	.333	.667	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
5-p	.000	.000	.400	.600	.000	.000	.000	.000	.000	.000	.000	.000
6-p	.000	.000	.000	.000	.455	.545	.000	.000	.000	.000	.000	.000
8-p	.000	.000	.000	.000	.000	.000	.455	.545	.000	.000	.000	.000
9-p	.000	.000	.000	.000	.000	.000	.000	.000	.400	.600	.000	.000
10-p	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.333	.667

The Matrix TA Split Into Two Parts

Again, the TA matrix shows the probability of ending in any given absorbing state having started in any given transient state. In this case, there is a unique starting state that we have by pure chance labeled *Start*. The probability of winning a game of craps is the combined probabilities of winning states marked 7, 11, 4−4, 5−5, 6−6, 8−8, 9−9, 10−10. The sums of those probabilities shown here is .493 whereas the correct probability is .492929.... The first part of the matrix shown in the upper part of the box shows the probabilities of winning or losing the game on the first throw, which can only happen from the *Start* state. The second part of the matrix shows that if your point is either 4 or 10, then your probability of winning is precisely 1/3. If your point is 5 or 9, your probability of winning is precisely 2/5, and if your point is 6 or 8, your probability of winning is precisely 5/11.

<u>Example</u>

The following example is of importance because it illustrates the application of the preceding techniques for the case of a Markov chain with transient and ergodic states that are not necessarily absorbing. Given the chain in **Figure 3**, we ask, what is the long probability term probability of being in state B? Note, that this problem is very sensitive to where we start. For example, if we start in state A, the long term probability of being in state B is 0. If on the other hand we start in state B or state C, the long term probability of being in state B is .5 because B−C is an ergodic sub-chain with both B and C having .5 probability. Let us suppose that our initial state vector is (0, .2, 0, .5, .3, 0). That is, we start in state B with probability .2, and in state D with probability .5, and state E with probability .3.

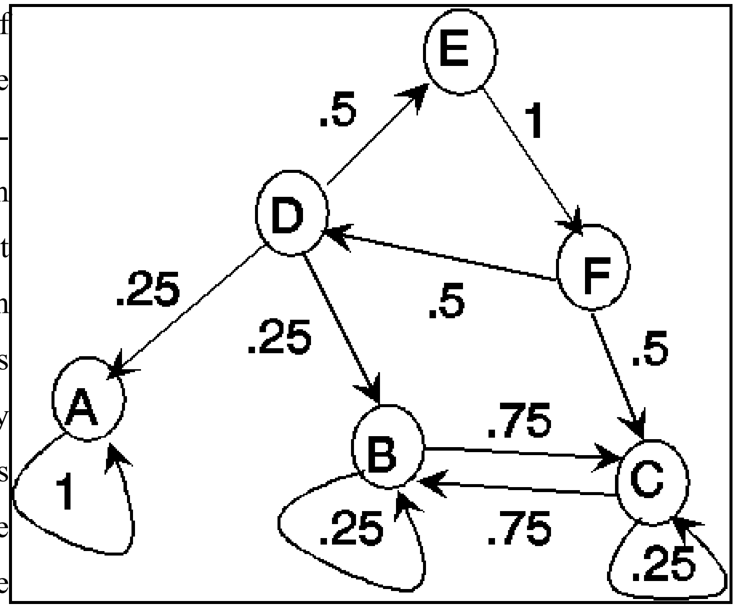

Figure 3 A Mixed Markov Chain

To further analyze the problem it is helpful to replace the chain of 9 with the Markov chain of **Figure 4**. Here we have replaced the ergodic nodes B and C with an absorbing state B'. The matrix analysis of this chain is given below.

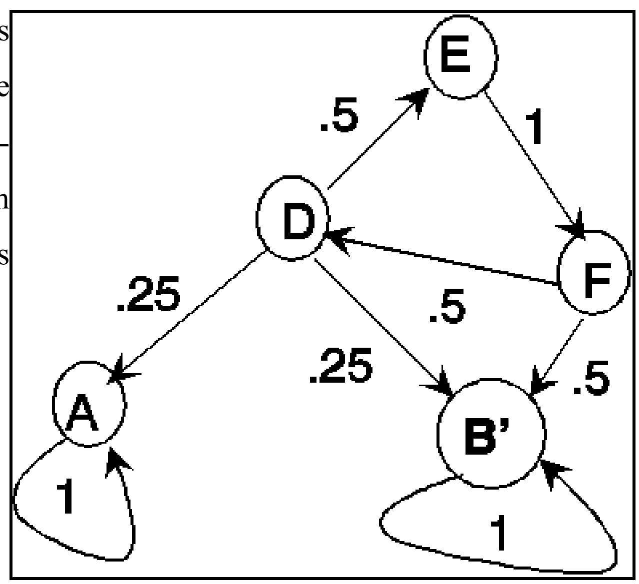

Figure 4 The Chain of **Figure 3** Reduced.

	D	E	F
D	1.333	.667	.667
E	.667	1.333	1.333
F	.667	.333	1.333

	A	B'
D	.333	.667
E	.167	.833
F	.167	.833

The Matrices TT and TA

The matrix TT gives the expected number of transitions from one transient state to another before being absorbed. It is the matrix, TA, on the right that is of interest to us. If we are in state D our probability of being absorbed in state B' is 2/3, but if we are in state E or F our probability of being absorbed in B' is 5/6. We have an initial probability vector of $P_0 = (.5, .3, 0)$ for states D, E, and F. Multiplying P_0 times TA we get:

$$P_0 \cdot TA = \begin{pmatrix} .5 & .3 & 0 \end{pmatrix} \begin{pmatrix} .333 & .667 \\ .167 & .833 \\ .167 & .833 \end{pmatrix} = \begin{pmatrix} .217 & .583 \end{pmatrix} = P_\infty$$

Note that neither P_0 nor P_∞ is a true probability vector since they each add to .8 and not 1. The remaining .2 is the probability of starting in state B which is part of the ergodic sub-chain given as B' in the reduced form of the chain. The vector P_∞ says that if one starts in state D with probability .5 and state E with probability .3 that the probability of ending in state A is .217 and the probability of ending in state B' (the ergodic sub-chain B−C) is .583. Hence the total probability of ending in state B' is .583 plus .2 (the probability of starting in B') which is .783. So for the starting vector (0, .2, 0, .5, .3, 0), the probability of ending in A is .217 and of ending in the sub-chain B−C is .783. That means there is a .783 probability that the chain will spend half of its time in sate B. The answer to our original question, *what is the long term probability of being in state B*, is .5·.783 = .391.

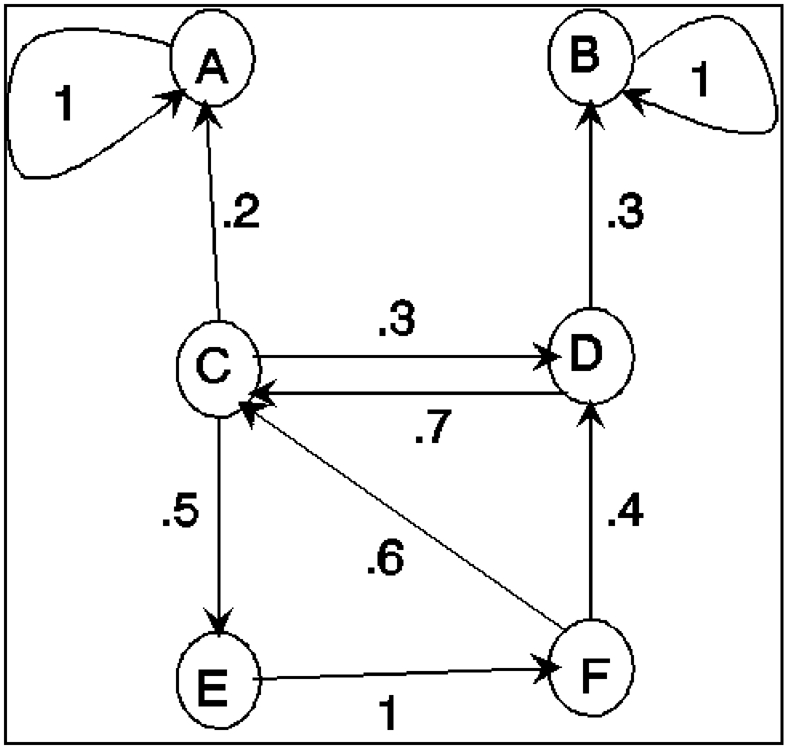

Figure 5 So, You Thought You Would Get Through This Section Without Homework

☐ **Exercise 1** Solve for the matrices TT and TA for the Markov chain of **Figure 5**.

☐ **Exercise 2** If you start in state E what is the expected number of trips to state D before being absorbed?

☐ **Exercise 3** If you start in state E, what are the probabilities of being absorbed in states A and B?

☐ **Exercise 4** Given the initial state vector P_0 = (.1, 0, .2, .4, .3, 0) what is the probability of being absorbed in state A?

1.

$$TT = \begin{bmatrix} 2.857 & 1.429 & 1.429 & 1.429 \\ 2 & 2 & 1 & 1 \\ 2.514 & 1.657 & 2.257 & 2.257 \\ 2.514 & 1.657 & 1.257 & 2.257 \end{bmatrix} \qquad TA = \begin{bmatrix} 0.571 & 0.429 \\ 0.4 & 0.6 \\ 0.503 & 0.497 \\ 0.503 & 0.497 \end{bmatrix}$$

2.

$$(0 \quad 0 \quad 1 \quad 0) \cdot TT = (\, 2.514 \quad 1.657 \quad 2.257 \quad 2.257 \,)$$

The expected number of trips to D after starting in E is 1.657

3.

$$(0 \quad 0 \quad 1 \quad 0) \cdot TA = (\, 0.503 \quad 0.497 \,)$$

4.

Start -------$(.1 \quad 0 \quad .2 \quad .4 \quad .3 \quad 0)$

$$(.2 \quad .4 \quad .3 \quad 0) \cdot TA = (\, 0.425 \quad 0.475 \,)$$

Probability of being absorbed in state A is .1 + .425 = .525