1 FTMC for countably infinite chains

1.1

To prove this, we only need to show $[F] + [A] + [I]$ implies $[PR] + [A] + [I]$.
To show this, we only need to show $[F] + [I]$ implies $[PR]$, i.e. a finite and irreducible markov chain is positive recurrent.

Firstly, for a finite markov chain, there exists at least one recurrent state. If not, all the states in this markov chain can only be visited finitely many times. But this finite markov chain only contains finite states, so when we visit infinitely many times this cannot be true. Then if the markov chain is also irreducible, since recurrence is a class property, the existence of a single recurrent state means this markov chain is recurrent.

Now we only need to show a finite irreducible recurrent markov chain is positive recurrent. Similar to our proof in lecture notes, we know that positive recurrence and null recurrence are both class properties. Then if the recurrent markov chain contains a null recurrent state, it is null recurrent itself. Since $\sum_{j \in S} P_{ij}^n = 1$ and there are only finite number of states, it is impossible that $\lim_{n \to \infty} P_{ij}^n = 0$ for all $j \in S$. Thus this markov chain cannot be null recurrent. Then it can only be positive recurrent.

To sum up, a finite irreducible markov chain must be positive recurrent. So the proposition in this problem about these two implications is correct.

1.2

Denote $P(i, i')P(j, j')$ by $Q((i, j), (i', j'))$. To prove $\Pr_{(i,j)}[T_{(k,k)} < \infty] = 1$, we only need to prove that Q is positive recurrent.
To prove this, we first show that Q has a stationary distribution. Note that P is positive recurrent and irreducible, from what we have proven in class, we know that P must have a unique stationary distribution. Denote the stationary distribution by π, then we know that the stationary distribution of Q is when it follows the distribution of π coordinate-wisely. So Q has a unique stationary distribution.
Then we show that Q is irreducible. For any i, j, i', j', we want to find a n such that $Q^n((i, j), (i', j')) = P^n(i, i')P^n(j, j') > 0$. Since P is aperiodic and irreducible, for any j, i' and sufficiently large n, $P^n(j, i') > 0$ always hold. So the n such that $Q^n((i, j), (i', j')) > 0$ always exists. Hence Q is irreducible.
Now we show that for any markov chain, irreducible provided, if the chain has a unique

stationary distribution, then it must be positive recurrent. Let $N_i(n)$ be the number of visits of state i in the first n steps. Since the markov chain is irreducible, the strong law of large numbers for markov chain shows that

$$P_j \left[\lim_{n \to \infty} \frac{N_i(n)}{n} = \frac{1}{\mathbb{E}_i(T_i)} \right] = 1$$

Denote the stationary distribution by π and let $X_0 \sim \pi$, then by the bounded convergence theorem we have

$$\mathbb{E}_{X_0 \sim \pi} \left[\frac{N_i(n)}{n} \right] = \sum_{t=1}^{n} \frac{\mathbb{E}_{X_0 \sim \pi} \left[1_{[X_t = i]} \right]}{n} = n \frac{\pi(i)}{n} = \frac{1}{\mathbb{E}_i(T_i)}$$

Since the chain is irreducible, $\pi(i) > 0$ for every i, thus $\mathbb{E}_i(T_i) < \infty$ for every i, i.e. the markov chain is positive recurrent.

Since Q is such a markov chain irreducible with a unique stationary distribution, it is positive recurrent. Then it follows that $\Pr_{(i,j)}[T_{(k,k)} < \infty] = 1$ for any $i, j, k \in \Omega$.

1.3

We already know from class that $[I] + [PR]$ implies $[S] + [U]$. So we only need to prove that the markov chain will converge if it is aperiodic.

We use total variance distance to describe the convergence. We want to show for any starting distribution μ_0, the total variance distance of μ_t and stationary distribution π, $D_{TV}(\mu_t, \pi)$ goes to 0 as t goes to ∞.

Follow the instructions of finite case from class, we will use two markov chains with the same transition matrix P. Suppose these two markov chains are $\{X_t\}$ and $\{Y_t\}$ where $Y_0 \sim \pi$ yet $X_0 \sim \mu_0$, an arbitrary distribution.

Now we construct a coupling ω_t of μ_t and π. The coupling is the same as finite case in class. Let $(X_t, Y_t) \sim \omega_t$, we construct ω_{t+1} by the following method. If $X_t = Y_t$ for some $t \geq 0$, then let $X_{t'} = Y_{t'}$ for all $t' > t$. If not, we let X_t and Y_t go to X_{t+1} and Y_{t+1} independently. The coupling lemma tells us that $D_{TV}(\mu_t, \pi) \leq \Pr(X_t \neq Y_t)$. And we have shown in class that with a finite markov chain

$$\lim_{t \to \infty} \Pr(X_t \neq Y_t) = \lim_{t \to \infty} \sum_{x_0, y_0} \mu_0(x_0) \pi(y_0) \Pr_{x_0, y_0} (X_t \neq Y_t) = 0$$

Now we consider the countably infinite case. For our coupling the transition function before $X_t = Y_t$ is what we have defined in problem 1.2, denoted by Q. From what we have proven in problem 1.2, $\Pr_{(i,j)}[T_{(k,k)} < \infty] = 1$ for any $i, j, k \in \Omega$. This means that the probability of X_t "meets" Y_t within finite time is 1, which is equivalent to saying that $\lim_{t \to \infty} D_{TV}(\mu_t, \pi) = 0$. So such a positive recurrent countably infinite markov chain will converge to its unique stationary distribution.

2 A Randomized Algorithm for 3-SAT

2.1

We use the same notations as defined in class. Similar to our analysis in class, if we repeat the random flipping operation for $2n^2$ times, the lower bound for

$$1 - \Pr\left[\exists t \in [0, 2n^2] \quad s.t. \quad Y_t = n\right]$$

is equal to

$$\Pr\left[T_{Y_0 \to n} > 2n^2\right] \le \frac{\mathbb{E}[T_{Y_0 \to n}]}{2n^2} \le \frac{1}{2}$$

Thus the probability that this algorithm is **not** correct after one iteration is $\frac{1}{2}$. But we will run our algorithm 50 times, if after any iteration our algorithm gives a correct assignment, then the algorithm is correct. If none of the 50 iterations says the CNF is satisfiable, then the probability of the CNF indeed being satisfiable is $(\frac{1}{2})^{50}$. Thus the probability of our algorithm is correct is $1 - (\frac{1}{2})^{50}$.

2.2

Without loss of generality, assume we choose clause $c = x \vee y \vee z$ at round t. Then we have $\sigma_t(x) = \sigma_t(y) = \sigma_t(z) = F$ where F indicates that those variables are false. And there are seven possible assignments of $\sigma(x), \sigma(y)$ and $\sigma(z)$, of which three are only one of x, y, z is T, three are exactly two of x, y, z are T, one is all of x, y, z are T.
In the first three cases $\Pr[X_{t+1} = X_t + 1] = \frac{1}{3}$ and $\Pr[X_{t+1} = X_t - 1] = \frac{2}{3}$.
In the middle trree cases $\Pr[X_{t+1} = X_t + 1] = \frac{2}{3}$ and $\Pr[X_{t+1} = X_t - 1] = \frac{1}{3}$.
In the last case $\Pr[X_{t+1} = X_t + 1] = 1$ and $\Pr[X_{t+1} = X_t - 1] = 0$.
Thus we have $\Pr[X_{t+1} = X_t + 1] \ge \frac{1}{3}$ and $\Pr[X_{t+1} = X_t - 1] \le \frac{2}{3}$.

2.3

We use the same notations in class. Like what we have done in class, we have the first hitting time

$$T_{i \to i+1} = \mathbb{1}[\mathcal{A}] + \mathbb{1}[\overline{\mathcal{A}}](1 + T_{i-1 \to i} + T_{i \to i+1})$$

where $\mathcal{A}$ represents the event that the first step is to the right.
Then by taking the expectations of both sides we have

$$\mathbb{E}[T_{i \to i+1}] = \frac{1}{3} + \frac{2}{3}(1 + \mathbb{E}[T_{i-1 \to i}] + \mathbb{E}[T_{i \to i+1}])$$

equivalently,

$$\mathbb{E}\left[T_{i\to i+1}\right] = 2\mathbb{E}\left[T_{i-1\to i}\right] + 3$$

Then by simple calculation we derive $\mathbb{E}\left[T_{i\to i+1}\right] = 2^{i+2} - 3$, by which we can further calculate that

$$\mathbb{E}\left[T_{i\to n}\right] = \sum_{k=i}^{n-1} \mathbb{E}\left[T_{k\to k+1}\right] = \sum_{k=i}^{n-1}(2^{k+2} - 3) = 2^{n+2} - 2^{i+2} - 3n + 3i \le 2^{n+2}$$

So the expectation of first hitting time of n from i is less than 2^{n+2}, which means that to reach a 0.99 correct probability, it is sufficient if we do the random flipping $400 \cdot 2^n$ times. Meanwhile it is impossible for us to finish this task within $O(2^n)$ times by our analysis, so we need to do the random flipping operation for $\Theta(2^n)$ times.

2.4

The probability $\Pr\left[\exists t \in [1, 3n] : X_t = n\right]$ expressed in human language is the probability of reaching $X_t = n$ within $3n$ steps with a starting point $X_0 = n - i$. We start at a situation that our initial random guess has $n - i$ variables correctly assigned and want to know the probability whether we can set other i variables properly within $3n$ steps.

Since we are trying to find a good lower bound for this probability, for $X_0 = n - i$, we can consider the initial $3i$ steps of the total $3n$ steps. If there are more than or equal to $2i$ steps in the $3i$ steps to the right, we can reach $X_t = n$ and hence we stop and get a correct assignment. Since we are finding a lower bound, we can only consider the case where exactly $2i$ steps are to the right and i steps are to the left. Now the probability can be expressed as

$$\binom{3i}{2i}\left(\frac{1}{3}\right)^{2i}\left(\frac{2}{3}\right)^{i} = \binom{3i}{i}\left(\frac{1}{3}\right)^{2i}\left(\frac{2}{3}\right)^{i}$$

By Stirling's formula, for $i \ge 2$, we have

$$\binom{3i}{i} \ge \frac{1}{\sqrt{5i}}\left(\frac{3^{3i}}{2^{2i}}\right)$$

Thus

$$\binom{3i}{i}\left(\frac{1}{3}\right)^{2i}\left(\frac{2}{3}\right)^{i} \ge \frac{1}{\sqrt{5i}}\left(\frac{3^{3i}}{2^{2i}}\right)\left(\frac{2^i}{3^{3i}}\right) = \frac{1}{\sqrt{5i}} \cdot \frac{1}{2^i}$$

So the lower bound for this probability is

$$\frac{1}{\sqrt{5i}} \cdot \frac{1}{2^i}$$

.

2.5

Based on the analysis of last problem, if we start with some σ_0 which is uniform at random from all 2^n assignments, then we need to add up the probabilities with $X_0 = n - i, i \in [n]$. The probability of $X_0 = n - i$ is simply

$$\binom{n}{i} \cdot \frac{1}{2^n}$$

Then by using the total probability rule, we can calculate the lower bound of the probability that the algorithm outputs a satisfying assignment.
The lower bound can be expressed by

$$\sum_{i=0}^{n} \binom{n}{i} \cdot \frac{1}{2^n} \binom{3i}{i} \left(\frac{1}{3}\right)^{2i} \left(\frac{2}{3}\right)^i \geq \frac{1}{\sqrt{5n}} \cdot \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \left(\frac{3^{3i}}{2^{2i}}\right) \left(\frac{2^i}{3^{3i}}\right) \tag{1}$$

$$= \frac{1}{\sqrt{5n}} \cdot \frac{1}{2^n} \sum_{i=0}^{n} \binom{n}{i} \cdot \frac{1}{2^i} \tag{2}$$

$$= \frac{1}{\sqrt{5n}} \cdot \frac{1}{2^n} (1 + \frac{1}{2})^n \tag{3}$$

$$= \frac{1}{\sqrt{5n}} \cdot \left(\frac{3}{4}\right)^n \tag{4}$$

Hence the probability that the algorithm outputs a satisfying assignment with n variables has such a lower bound.

2.6

Our strategy is to run the algorithm on 3-SAT questions for several times until it can find a solution with probability 0.99.
In previous problem we have calculated the probability for a single running for $3n$ rounds, which is $\frac{1}{\sqrt{5n}} \left(\frac{3}{4}\right)^n$. We use p to denote this probability, then if we run the algorithm for α times, the probability of finding a solution(if it exists) is

$$1 - (1 - p)^\alpha$$

To guarantee a probability larger than 0.99, we let

$$1 - (1 - p)^\alpha \geq 0.99$$

and get

$$\alpha \geq \frac{\ln 0.01}{\ln(1 - p)}$$

which is approximately

$$\frac{\ln 100}{p}$$

for n large enough.

So if we let

$$\alpha \geq \ln 100 \sqrt{5} n^{\frac{1}{2}} \left(\frac{4}{3}\right)^n$$

we can guarantee our algorithm can find a solution with probability 0.99.

Since we are going to run $3n$ rounds each time, the total running time is

$$cn^{\frac{3}{2}} \left(\frac{4}{3}\right)^n$$

and hence the smallest c for the required running time is $\frac{4}{3}$.

3   Acknowledgements

Q1.1

The idea of proving a finite irreducible markov chain is from lecture 5 of STAT253/317, University of Chicago.

Q1.2-1.3

Actually from lecture notes previous year.

Q2.4-2.6

I found online the paper by Uwe Schoning in 1999 where he described how to implement such an algorithm and do mathematical analysis.
I think the his illustration and analysis helped me a lot when I's solving this problem.