Q 2:

- Bellman eq for m:

$$Q^*_M = \sum_{s'} P(s'|s,a)\left[R(s,a,s') + \gamma \max_{a'} Q^*_M(s',a')\right]$$

- Bellman eq for m':

$$Q^*_M = \sum_{s'} P(s'|s,a)\left[R'(s,a,s') + \gamma \max_{a'} Q^*_{M'}(s',a')\right]$$

Plug in R' as R

$$= \sum_{s'} P(s'|s,a)\left[R(s,a,s') + \phi(s) - \gamma\,\phi(s') + \gamma \max_{a'} Q^*_{M'}(s',a')\right]$$

$$= \sum_{s'} P(s|s,a) R(s,a,s') + \sum_{s'} P(s'|s,a)\,\phi(s) - \gamma \sum_{s'} P(s'|s,a)\,\phi(s') + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q^*_{M'}(s',a')$$

$\phi(s)$ is independent of s'

$$= \sum_{s'} P(s|s,a) R(s,a,s') + \phi(s) \sum_{s'} P(s'|s,a) - \gamma \sum_{s'} P(s'|s,a)\,\phi(s') + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q^*_{M'}(s',a')$$

$\sum_{s'} p(s'|s,a) = 1$

$$= \sum_{s'} P(s|s,a) R(s,a,s') + \phi(s) \cdot 1 - \gamma \sum_{s'} P(s'|s,a)\,\phi(s') + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q^*_{M'}(s',a')$$

$$\Rightarrow \quad Q^*_{M'}(s,a) = Q^*_M(s,a) + \underbrace{\phi(s) - \gamma \sum_{s'} P(s'|s,a)\,\phi(s')}$$

Since this term is a state dependent and independent of a, it does not affect on the slection of the optimal policy.

$\Rightarrow$ The optimal policy for m and m' are identical //

Q 4:

1)
- from the starting position $(b,1)$ the legal moves are:

$(b,1) \to (a,3)$
$(c,3)$
$(d,2)$

- to return to $(b,1)$ after 2 steps the knigt must move to position which he can retrurn to $(b,1)$, bossible moves are:

$(a,3) \to (b,1)$
$(c,3) \to (b,1)$
$(d,2) \to (b,1)$

- So, initially the knight has 3 legal moves
  - from each it can return to $(b,1)$

$\Rightarrow P = 1/3 \cdot \frac{1}{8} \cdot 3 = 1/8$ //

2)- the markov chain is irreducible since its possible to get from any state to any other,
for the knigt it can evetually get to any squere from any starting position.

- the markov chain is periodic since the knigt can return to any state every "pair" steps (מספר זוגי של צעדים) , thus the biggest dividor is 2 , thus periodic with $T=2$

3) the mean recurrence time T_i to a state i is the reciprocal of the stationary distribution π_i for that state

Since the knight spends an equal amount of time on each square :

$$\pi_i = 1/64 \quad \forall i$$

$$\Rightarrow$$

$$T_{corner} = \frac{1}{\pi_{corner}} = 64$$