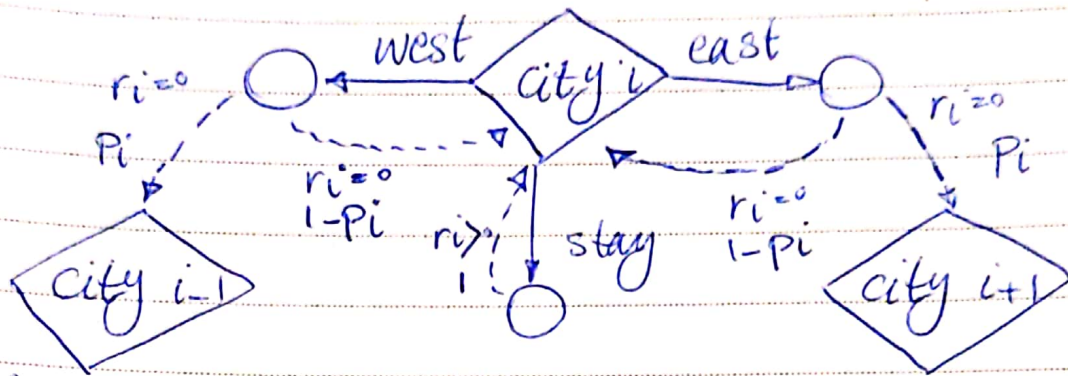


تمرین کتبی سری دوازدهم خوش منوی

سوال اول الف



fixed policy: $U_{stay}(i) = \sum_s T(s, stay(s), s') [R(s, stay(s), s') + \gamma U_{stay}(s')]$

$$U_0 = 0 \rightarrow T(s, stay(s), s') = P_i = 1$$

$$U_{stay,1}(i) = 1 \times (1) = 1, U_{stay,2}(i) = 1 \times (1 + 0.8) = 1.8$$

$$U_{stay,3}(i) = 1 \times (1 + 0.8 \times 1.8) = 2.44 \rightarrow \lim = 2 \rightarrow U_{stay}(i) = 2$$

($U_{stay}(i) = 2$)

	$Q(1, stay)$	$Q(1, east)$	$Q(2, west)$	$Q(2, stay)$
1, stay, 1, 1	2	0	0	0
1, east, 0, 2	2	0	0	0
2, stay, 1, 2	2	0	0	2
2, west, 0, 1	2	0	1.8	2
2, stay, 1, 1	2.8	0	1.8	2