Using the definition of a function and RUMAN equalion we oftain the following Q#18, Q = Em [ 3 ) KREHKHA 186 8, Ax = a ] = = Eq CRex1 + 3 7 1 Rex1 1 86 = 3 Ax = 9 3 = # B [ R + + 1 S = \$, A = a ] = B + ( 9 a ) ; ET Z X KREIK+1 SE= & AE= Q ] = = FALY 2 YKREHKA1 1 SEN = S' ACHI F 9' ] = FALY QA 15' Q ] ] > 2 (5, a) = E15, a) n P(ils, a) 1 / (5, a) + / (15, a) ) 2. Q \* /5, a) = max q = (5, a) = mox (E(5, a)) ~ P(1/5, a) + / a (5, a') } = E(5), a) ~ 5 \* ( (/5, a) + / max a (/5, a)) = = B15'a)~ 4.18a) [ [15a) + 1 max Q#15'a')

Plausable obsertive minimizes the difference between current a-function and the optimize Q-function e(0) = = 11 Q\*18 a,0) - Q(9, a, 0) 112 Q\* (5 a, 6) = Esin (15 a) [ (15 a) + 1 max (2\*15, a) ] = ESTA (18,0) & 18,0) & max wax (215,0) > (10) = First (189) 11 [+ Trac max (215) a) - (218,99) 118