Christian (Joshua A	berto	COM 221				
Exercise 6					Step		
of the second					Target		
1. Draw the	Q. table str	icture) = 1.5	JE 13-0+0	0(827)=		
SARSA							
State	Up	Down	Left	Right	Episode 2904		
0,0	-0.05	-0.05	0	6	(tep (: (0,0		
0,1	. 0	Ó	1000	=09 F	1,0) : sport		
0, 2	0	[0.7]	05	0	(S.O) 18 got?		
1,0	0	t0:05	Col 0 5 (6)	-0.05	Stopy: (ip)		
	0	1000	6	-0.05			
1,2	O	(1)	6	-10 85	Step 1 3		
2,0	Ø	0	6)0=	= 0 + 1.0-	E PART		
2,1	0	0 80.0	0 . (1-)@.0-0=	(900) 8		
2,2	2	0	6	δ	15002		
				1 = 5+	= +====		
Episode Pas	th		2.0 = (e-	-1 +2.0+0=	(810)0=		
Step 1: (0,0) -> 17 -	ا (١٥٥) [veward -0.	·()	Step 3:		
Step 2: (1,0) -7 R -	7 (1,1) [reword -0-1](ae) + 1.0	-= torr		
Step 3: (1,1) -7 R=	(1,2)	rehard -0.	(g) 2.0+0=	(0.50)		
Step 4: (1,2) -> p-	(2,2) [used +3		4 32		
	Programme Progra						
Step 1 =		2.25	F = (21-	r) 3.0 + 6.4	7= 始;) 20		
	-01+0:						
Q (0,2,D)	= 0 +0.5 ((-01-0)	=-0.05				
5kp 2:							
target:	2 -0.1 to z	-0.1					
	= 0 +0.5(0.05	Alternative Control			
Step 3:					7		
torget:	0.1 + 0 =	-0.1					
$Q(1,1,R) = 0 + 0.5(-0.1-0)^{2} - 0.05$							

```
Step 4:
Target
 Q(4,2,0) = 0 + 0.5 (3-6) = 1.5
Episode 2 Path
 Step 1: (0,0) - R-7
                      (0,1)
                             reword -0-1
 Step 2: (0,1) -7 R-7 (0,2) record + 1
 Step 3: (0,2) -> D-> (1,2) reward -0.1
 Step4: (1,27 - 107)
                      (2,2) havard +3
Step 1:
   Target = -0.1 + 6 = -0.1
   Q (0,0,R) = 0+ 6.5 (-0.1-0) = -0.05
Step 2:
   Target = 1 +0 = 1
   Q(0,1,R) = 0 +0.5(1-0) = 0.5
Step 3:
                   [10- paras] (01) L
   Target =-0.1+(0.5) = (0.4)
   Q(0,2,0) = 0 + 0.5 (0.4-0) = 07 (5)
 Step 4:
   Farget = 3 + 02
   Q(43) = 1.5 + 6.5(3-15) = $2.25
```

a learning	9					Ep 2 Posts		
State	νp	Down	Left	Right	DEFF	(0,0) · t will		
0,0	-0.05	-0.05	6	020 m	FAR	(Que 2: (0.1)		
0,1	0	2	1-0-0 mis	161)	101	Sty 3 = (0,2)		
0,2	6	0.7	E+ 0 +3	(2,0)	F-0 P	(1,2) - (xy)		
1,0	0	0	O	_ 0.05				
1,1	0	6	0	-0.05		Moste		
1,2	0	225	0	0	= 04	o s tomot		
20	0	0	:. - 005	(0)	77.0 × 0	R(O,OR)		
2,1	0	0	0	0		1402:		
2,2	0	6	0	Ò	15 0+	Tomet > 1		
			70.8-	= (0-10	7.04	0= /21.010		
Episode Pal	th	AC				18 m		
Step 1: (D-7 (40)	rward	0.10.1 =	+ (0.5)	0 15moT		
Sup 2: (R 7 (1,1)	mard -	0.(4.0) 2	0 + 6.	= (900)		
Step 3:		R - (1,2)	4-010-	١.م.		: P 9-K		
Step 4:	the state of the s	075 (2,2) rward.	+3	(C.S.)	1 g scent		
Step 1								
Targes	=-0.1 +0	= 0.1						
Q(1,01	0) = 0+0.5	(-0.1-0)	=-0.05					
Step 2								
Target	= -0.1 +0	= -0-1			1000			
Q(1,1,R) = 0+0.5(-0.1-0) = 0.05								
(tep 3								
T 1	= -0-1-	to = 0.1				•		
Q(1,2,R) ~0 +0.5 (-0.1-0) = 0.05								
Carol								
Q(12)	p) ~ O+	-0.5 (3-0)	= 1.5			- Process		

```
Ep 2 Path
Step 1: (0,0) -2 R-20, 1) revovor -0.1
Step 2: (0.1) 7 12 7 (0,2) mand +1
 Sty 3: (0,2) -7 D = (1,2) neward - 0-1
 Step 11: (1,2) -7 D-7 (2,2) mare +3
 Step1:
  Target = 0.1 to = -0.1
 Q(0,0,R) = 0+0.5(-0.1-0) = -6.05
 Step 2:
  Target = 1+0 =1
  Q(0,1,R) = 0+0.5 (-0.1-0) =-6.05
 Step 3:
 Target =-0.14 (0.5) = 1.40 Anoun (01)
  Q(0,2,Q) = 0+6.5 (0.4-0) = 0.7
 Step 4:
   Target Q((1,2,0) = 1.5 +0.5(3-1.5) = 2.25
```

0.0-

70.0 = (0-115-)2010 = 10/11

20 (= (0-1.0-) 2.04.0 - 10 19

JI - Part Fact

10 - 1AG