Sample exercises

$WA^* = g + w \cdot h$

$A^* = g + h$.

$UCS = g$

Best first search $= h$.

1 (a) BFS: queue FIFO.

DFS: stack: LIFO.

UCS: priority queue ordered by g.

minimum cost g, maximum priority

(b)

I: S_1

goal: S_4.

~~cost~~	h
S_1	12
S_2	9
S_3	5
S_4	0

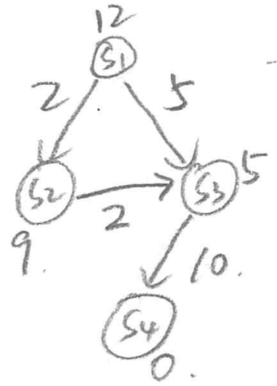
~~cost~~

admissible $h_{(s)} \le h^*$

~~admissible~~ consistant

$A(s) \le c(s') + h(s')$

$\Leftarrow s \to s'$



iteration	open	closed
0	$S_1(12)$	Φ
1	$S_2(11)$ $S_3(10)$	$S_1(12)$
2	$S_4(15)$ $S_2(11)$	$S_1(12)$ $S_3(10)$
reopen 3	$S_3(7)$ $S_4(15)$	$S_1(12)$ $S_2(11)$ $S_3(10)$
4	$S_4(15)$	S_1 S_2 S_3
5	Φ	$S_4 \to$ stop

$S_1 \to S_2 \to S_3 \to S_4$

[illegible] reopen

2

~~B~~ Rooms $= \{A, B\}$, Objects $= \{O_1, O_2\}$

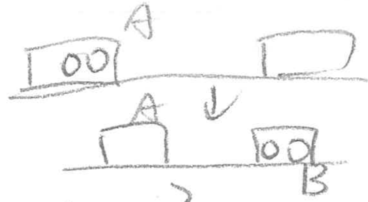
$F = \{$robot(r), holding(o), ~~object~~ at(r, o), free $|$ r $\in$ Rooms, o $\in$ Objects$\}$

A = {pick(O, R):

pre: free, ~~object~~ at(R, O), robot(R)

add: holding(O),

del: free, at(R, O). | O $\in$ Objects, R $\in$ Rooms}



A = A $\cup$ {move(r_1, r_2):

pre: robot(r_1)

add: robot(r_2)

del: robot(r_1) | r_1, r_2 $\in$ Rooms}

I = {at(A, O_1), at(A, O_2), robot(A), free}

G = {at(B, O_1), at(B, O_2)}

A = A $\cup$ {drop(O, R):

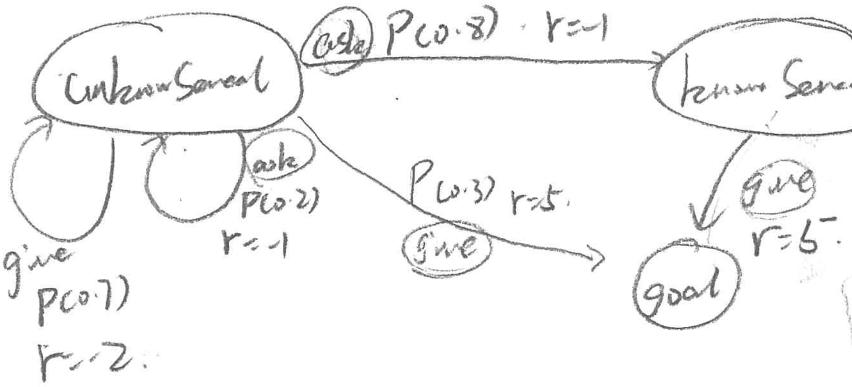
pre: holding(O), robot(R)

add: at(R, O), free.

del: holding(O) | O $\in$ Objects, R $\in$ Rooms}

Q $\boxed{3}$

$Q(s,a) = \sum P_a(s'/a)\,[r(s,a,s') + \gamma V(s')$



Iteration	1	2	3
V(known)	0	5	5
V(unknown)	0	0.1	2.618

(a). $\gamma = 0.9$

$\boxed{I2}$

$V_{know} = \max Q(\text{know}, \text{give}) = 1\times[5+0.9\times0] = 5$

$V_{unknow} = \max \begin{cases} Q(\text{unknow}, \text{ask}) = -1 \\ Q(\text{unknow}, \text{give}) = 0.1 \end{cases}$

$Q(\text{unknow}, \text{ask}) = 0.8\times[-1+0.9\times0] \; \underset{\cancel{-0.8}}{} + 0.2[-1+0.9\times0] = -1$

$\cancel{-0.8}$ $\quad -0.2$

$Q(\text{unknow}, \text{give}) = 0.3\times[5+0.9\times0] + 0.7\times[-2+0.9\times0] = 1.5 - 1.4 = \underline{\underline{0.1}}$

$\boxed{I3}$: $Q(\text{know}, \text{give}) = 1\times[5+0.9\times0] = 5$

$Q(\text{unknow}, \underline{\underline{\text{ask}}}) = 0.8\times[-1+0.9\times5] = 3.5\times0.8 = 2.8$

$\quad\quad 3.5$

$\quad + 0.2\times[-1+0.9\times0.1] = -0.182 \quad = \boxed{2.618}$

$\quad 0.09 \quad -0.91$

$$\begin{array}{r} 3.5 \\ \times\, 0.8 \\ \hline 2.80 \end{array} \quad \begin{array}{r} 2.8 \\ -\,0.182 \\ \hline 2.618 \end{array} \quad \begin{array}{r} -0.91 \\ \times\, 0.2 \\ \hline -0.182 \end{array}$$

$Q(\text{unknow}, \text{give}) = 0.3\times[5+0.9\times0] + 0.7\times[-2+0.9\times0.1]$

$\quad 0.09$

$1.5 + (-1.337)$

$= \boxed{0.163}$

$$\begin{array}{r} -1.91 \\ \times\, 0.7 \\ \hline 1.337 \end{array} \quad \begin{array}{r} 1.5 \\ -\,1.337 \\ \hline 0.163 \end{array}$$

(b)

$\boxed{4}$ $x = 0.9$ $\partial = 0.5$ $Q(\cancel{a}s,a) = Q(s,a) + \partial [r + \gamma \max \cancel{Q} Q(s',a') - Q(s,$

(a) 1-step Q-learning:

$Q(s,a) = 1.9 + 0.5 \times [-2 + 0.9 \times \cancel{\max Q(\text{unknown})} \overset{7.2}{} - 1.9]$

(unknown give

6.48 2.58

$= 1.9 + 0.\cancel{5 \times 2.58} = \boxed{3.19}$

1.29

1-step-SARSA

$Q(s,a) = 1.9 + 0.5 \times [-2 + 0.\cancel{9 \times 1.9} - 1.9]$

$= \textcircled{0.805}$ -2.19 1.71 -1.095

(b) $Q(s,a) = Q(s,a) + \partial [G^2_{(t)} + \gamma^2 Q(s,a) - Q(s,a)]$

$= 1.9 + 0.5[-3.8 + 0.\cancel{81 \times 1.9 - 1.9}]$

1.539 -0.361

$= \boxed{-0.1805}$ -4.161 -2.0805

$G^2_{(t)} = -2 + 0.9 \times (-2) = -2 - 1.8 = -3.8$

(c) 用 use SARSA safe policy

$\partial \Rightarrow$ lower

$\boxed{Q6}$ M is dominated strategy. discard

player 2: y, $1-y$ ($\frac{1}{2}$, $\frac{1}{2}$)

player 1: x ($\frac{2}{5}$), $1-x$ ($\frac{3}{5}$)

	L	R
U	2,4	1,1
D	3,2	0,4

$E(U) = 2y + 1 - y = y + 1$

$E(D) = 3y + 0 = 3y$

$y + 1 = 3y$

$y = \frac{1}{2}$

$E(L) = 4x + 2(1-x) = 2x + 2$

$E(R) = x + 4(1-x) = -3x + 4$

$E(L) = E(R)$ $2x + 2 = -3x + 4$

$5x = 2$

$x = \frac{2}{5}$

Mixed Nash eqb:

$\{$(player 1: $\frac{2}{5}$ U, $\frac{3}{5}$ D),

(player 2: $\frac{1}{2}$ L, $\frac{1}{2}$ R) $\}$

$7.2 \times 0.9 = 6.48$

$4.48 - 1.9 = 2.58$

$2.58 \div 2 = 1.29$

$1.9 + 1.29 = 3.19$

$1.9 \times 0.9 = 1.71$

$1.9 - 0.29 = -2.19$

$2.19 \div 2 = 1.095$

$1.9 - 1.095 = 0.805$

1.9×0.81: 19, 152, 1.539

$1.9 - 1.539 = -0.361$

$3.8 + 0.361 = -4.161$

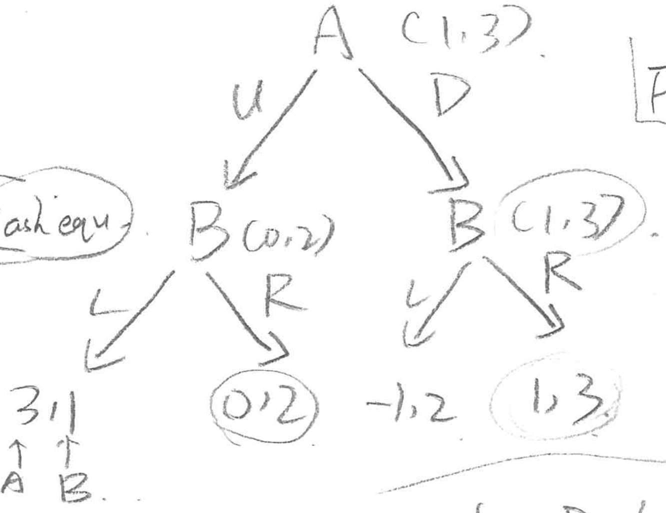
$4.161 \div 2 = 2.0805$

$2.0805 - 1.9 = 0.1805$

(b).

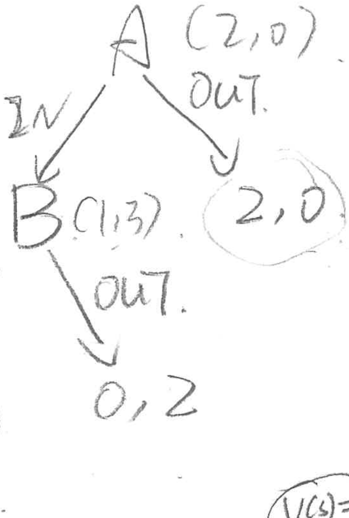
		B	
		L	R
A	U	3,1	0,2
	D	-1,2	1,3

Nash equ.

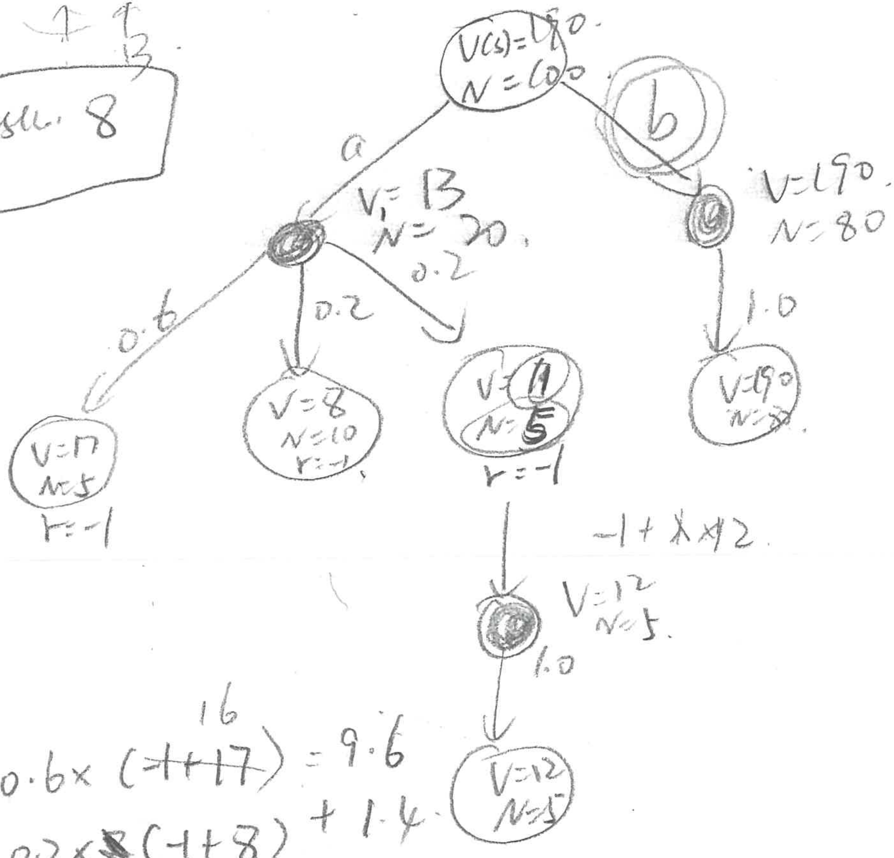


For A: D. B: R

A out. B. forced end game



question 8



$V_1 = 0.6\times(\overset{16}{\cancel{-1+17}}) = 9.6$

$+ 0.2\times\cancel{8}(-1+8) + 1.4$

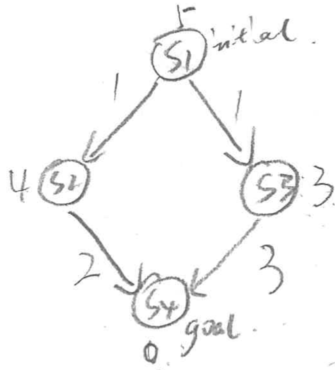
$+ 0.2\times(-1+11) + 2$

b.

$$\begin{array}{r} 16 \\ \times 0.6 \\ \hline 9.6 \end{array}$$

2017 Sample

a.



A* not find optimal ⇒ h not admissible

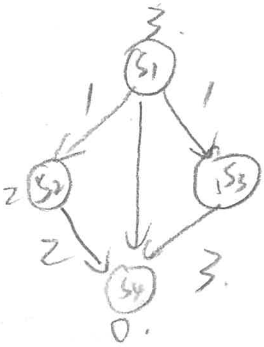
	h(s)
S1	5
S2	4
S3	3
S4	0

	c(s)
S1→S2	1
S1→S3	1
S2→S4	2
S3→S4	3

	open	closed
iteration 0	S1(5)	ϕ
i 1	S2(5) S3(4) now	S1(5)
i 2	S2(5) S4(4)	S1(5), S3(4)
i 3	S2(5)	S1(5), S3(4) S4(4) ✓ stop

S1→S3→S4 not optimal

b

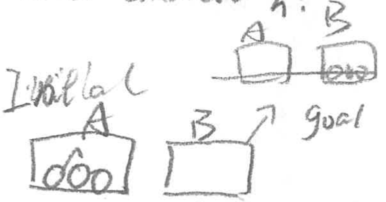


	h*(s)
S1	3.
S2	2
S3	3.
S4	0.

A* will expand S1→S2→S4.

hill-climbing will chose smallest h.

S1→S4



2. STRIPS P = <F, O, I, G>, Locations = {A, B}, ~~Rooms = {A, B}~~ }. Balls = {b1, b2, b3}

Atoms F = { robot(l), ball(b, l), holding(b). free | l ∈ locations, b ∈ Balls)

operations (actions) A = { move(~~from~~ x, ~~to~~ y):
pre = ~~holding~~ robot(x)
add = robot(y)
del = robot(x) | x, y ∈ Locations }

A = A ∪ { pickup(from, ball):
pre: free, robot(from), ball(ball, from)
add: holding(ball)
del: free, ball(ball, from), | from ∈ Locations, ball ∈ Balls }

A = A ∪ { putdown(from, ball):
pre: holding(ball), robot(from)
add: ball(ball, from), free
del: holding(ball) | from ∈ Locations, ball ∈ Balls }

I = { free, ball(b, A), robot(A) | b ∈ Balls }

G = { ball(b, B) | b ∈ Balls }

(b) Initial state goal state

~~robot(A), robot(B)~~, ball(b1, A), ball(b1, B), ball(b2, A) ball(b2, B) b1

iteration	robot(A)	robot(B)	ball(b1,A)	ball(b1,B)	ball(b2,A)	ball(b2,B)	ball(b3,A)	ball(b3,B)	holding(b1)	h(b2)	h(b3)	free
0	0	∞	0	∞	0	∞	0	∞	∞	∞	∞	0
1	0	1	0	∞	0	∞	0	∞	1	1	1	0.
2	0	1	0	3/2	0	3/2	0	3/2	1	1	1	0.
				add max		add max		add max				

move(from, to)

pickup(from, ball)

putdown(from, ball).

robot(B) = move(A, B)
↳ pre : robot(A).

ball(b1, B) = putdown(B, b1)
↳ pre = holding(b1), robot(B)

holding(b1) = pickup(A, b1)
↳ pre = free, robot(A), ball(b1, A)

h(add) = 3+3+3=9

h(max) = 2.

b3 doesn't change for h^{add} or h^{max}

I = {free, ball(b, A), robot(A) | b∈balls}.

G = {ball(b, B) | b∈Balls}.

bs(ball(b1, B)) = putdown(B, b1)
↳ pre = holding(b1), robot(B)

bs(holding(b1)) = pickup(A, b1)
↳ pre = {free, robot(A), ball(b1, A)
(supported by I)

bs(robot(B)) = move(A, B)
↳ pre : robot(A) supported by I.

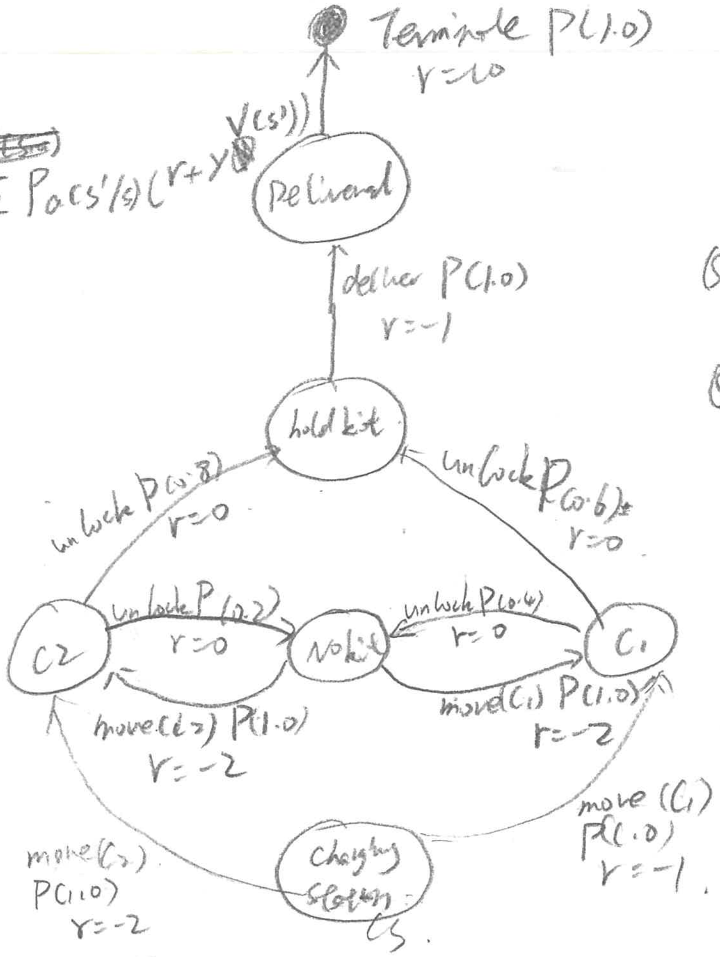
bs(ball(b2, B)) = putdown(B, b2}
↳ pre = holding(b2), robot(B).

relaxed plan = {pickup(A, b1), pickup(A, b2), pickup(A, b3), move(A, B)
putdown(B, b1), putdown(B, b2), putdown(B, b3)}.

h^{ff} = 7

3

$Q(s,a) = \cancel{Q(s,a)}$
$\sum P_a(s'|s)(r+\gamma V(s')))$



(1). $\gamma = 1.0$.

$$V(CS) = \max \begin{cases} Q(CS, move(C_1)) \\ Q(CS, move(C2)) \end{cases}$$

$$Q(CS, move(C_1)) = 1 \times (-1 + 1 \times 5.4) = 4.4$$

$$Q(CS, move(C2)) = 1 \times (-2 + 1 \times 7.2) = 5.2$$

$$\boxed{V(CS) = 5.2}$$

$\therefore \Rightarrow$ move (C2)

(2) $V(C_1) = \max (Q(C_1, unlock))$

$$= Q(C_1, unlock)$$
$$= P(No kit | C_1)(0 + 1 \times 0) + 0.6 \times (0 + 1 \times 9) = 5.4$$

$$\boxed{V(C_1) = 5.4}$$

$$V(No kit) = \max \begin{cases} Q(move(C1)) \\ Q(move(C2)) \end{cases}$$

$$Q(No kit, move(C_1)) = 1 \times [-2 + 1 \times 5.4] = 3.4$$

$$Q(No kit, move(C2)) = 1 \times [-2 + 1 \times 7.2] = 5.2$$

$$\boxed{V(No kit) = 5.2}$$

$$V(C2) = Q(C2, unlock)$$
$$= 0.2 \times [0 + 1 \times 0] + 0.8 \times [0 + 1 \times 9] = 7.2$$

$$\boxed{V(C2) = 7.2}$$

4

Robot B

Robot A	R1	R2	R3
R1	4,4	5,0	5,5
R2	5,5	5.0	5,5
R3	0,5	0,0	0,5

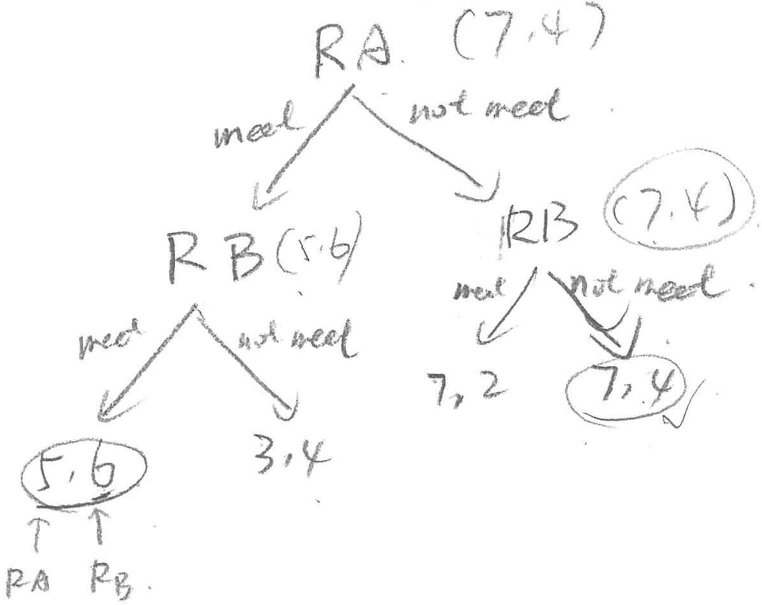
check all pure strategy equilibria

Robot B (L)

Robot A (K,M)	meet	Not meet
Meet	5,6	3,4
not Meet	7,2	7,4

Nash equilibrium (Robot A: not Meet, Robot B: not meet

higher payoff


RA (7,4)
meet
not meet
RB (5,6)
RB (7,4)
meet
not meet
meet
not meet
5,6
3,4
7,2
7,4
RA
RB