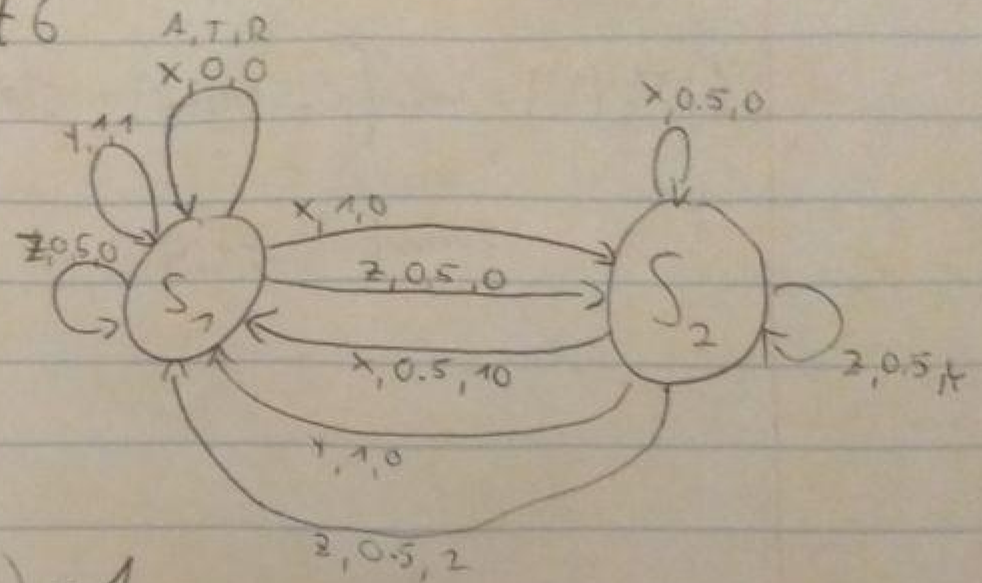
CSE415 - Assignment 6

1)



a) $V_0(S_1) = (1\cdot 0 + 0\cdot 0)$
$+ (1\cdot 1 + 0\cdot 0)$
$+ (0.5\cdot 0 + 0.5\cdot 0) = \underline{\underline{1}}$

b) $V_0(S_2) = \max(5, 0, 3) = \underline{\underline{5}}$

c) $V_1(S_1) = \max(5, 1, 2.5) = \underline{\underline{5}}$

~~d) $V_1(S_2) = \max(0.5\cdot(10+5), 1, 2$~~

~~d) $V_1(S_2) = \max(0.5\cdot(10+1), 1, (1+0.5)+2+2.5)$~~
~~$= \max(5.5, 1, 6) = 6$~~

~~e) $V_2(S_1) = \max(6, 1+5$~~

d) $V_1(S_2) = \max(0.5(10+1)+0.5(5), 1, 6) = \max((5.5+2.5), 1, 6)$

$= \underline{\underline{8}}$

e) $V_2(S_1) = \max(8, 1+5, 4+2.5) = \underline{\underline{8}}$

f) $V_2(S_2) = \max\left(\overset{7{,}5}{0.5(10+5)} + \overset{4}{0.5(8)},\ 5,\ \overset{3.5}{0.5(2+5)} + \overset{6}{0.5(4+8)}\right)$

$= \max\left(11.5,\ 5,\ 9.5\right) = \underline{\underline{11.5}}$

2a)

[A, E, C, 2]	$Q(A,E) = \underline{1}$
[C, E, F, 2]	$Q(C,E) = \underline{1}$
[F, S, G, 8]	$Q(F,S) = \underline{4}$
[C, S, P, -2]	$Q(C,S) = \underline{-1}$
[D, E, G, 8]	$Q(D,E) = \underline{4}$
[C, S, F, 2]	$Q(C,S) = 0.5 \cdot (-1) + 0.5 (2+4) = \underline{2{,}5}$
[C, E, D, -2]	$Q(C,E) = 0.5 \cdot 1 + 0.5 (4 + -2) = \underline{1.5}$

b) i)

$T(A,E,C) = 1/1 = \underline{1}$	$R(*,*,C) = \underline{2}$
$T(C,E,F) = \underline{1/2}$	$R(*,*,D) = \underline{-2}$
$T(C,E,D) = \underline{1/2}$	$R(*,*,F) = \underline{2}$
$T(C,S,F) = \underline{1/2}$	$R(*,*,G) = \underline{8}$
$T(C,S,D) = \underline{1/2}$	
$T(D,E,G) = \underline{1}$	
$T(F,S,G) = \underline{1}$	

ii) $V^*(F) = \underline{8}$ $\quad V^*(C) = 0{,}5 \cdot (8+2) + 0{,}5 (8-2) = 5 + 3 = \underline{\underline{8}}$

$V^*(D) = \underline{8}$ $\quad V^*(A) = 2 + 8 = \underline{\underline{10}}$

c) i) model-free

ii) model-based

~~2c) iii) Reinforcement learning:~~

2c) iii) Model-free:

A model-free algorithm does not need to learn T and R explicitly and is therefore often faster in learning the policy.

Model-based:

A model-based algorithm learns the whole MDP (S, A, T, R) and can therefore guarantee to find the optimal policy, while a model-free algorithm might converge to a sort of local minimum due to the exploration vs. exploitation tradeoff.

3)

		X: T	X: F	
Y	T	0,12	0,28	0,4
	F	0,18	0,42	0,6
		0,3	0,7	

a)
$$P(x) = P(x,y) + P(x,\bar{y})$$
$$= 0{,}12 + 0{,}18 = \underline{\underline{0{,}3}}$$

b) $P(y) = 0{,}12 + 0{,}28 = \underline{\underline{0{,}4}}$

c) $P(x) \cdot P(y) = 0.3 \cdot 0{,}4 = \underline{\underline{0{,}12}}$

d) Yes! Because $P(x) \cdot P(y) = P(x,y) := P(x|y) \cdot P(y)$
$$= P(y|x) \cdot P(x)$$

$$\Rightarrow P(x) = P(x|y)$$

for the numbers above. $P(y) = P(y|x)$

e) In general: (Assume $P(y) > 0$)

$$P(x|y) \cdot P(y) = P(x,y)$$
$$\Leftrightarrow P(x|y) = \frac{P(x,y)}{P(y)} = \frac{P(y|x) \cdot P(x)}{P(y|x) + P(y|\neg x)}$$

$$\Leftrightarrow P(x|y) = \frac{P(y|x) \cdot P(x)}{P(y|x) + P(y|\neg x)}$$

f) It doesn't matter when to buy Tesla stocks, since $X \perp\!\!\!\perp Y$. ~~Therefore on rainy and not rainy days the~~

4)

a) $P(A,B,C,D) = P(D \mid C) \, P(C \mid A,B) \, P(A) \, P(B)$

b)

		X_t	
		R	F
X_{t+1}	R	0,6	0,8
	F	0,4	0,2

5) a)

$$P(X_{t+1}=R \mid X_t=F)=0{,}8$$

b) $P(X_{t+2}=R \mid X_t=F)$

$$= 0{,}8 \cdot P(X_{t+1}=R \mid X_t=R) + 0{,}2 \cdot P(X_{t+1}=R \mid X_t=F)$$

$$= 0{,}8\, 0{,}6 + 0.2 \cdot 0.8 = 0{,}48 + 0{,}16 = 0{,}64$$

c) $P_\infty(R) = P_\infty(R) \cdot 0{,}6 + P_\infty(F) \cdot 0{,}8$

$$0{,}4\, P_\infty(R) = 0{,}8\, P_\infty(F)$$

$$P_\infty(R) = 2 P_\infty(F) \quad , \quad P_\infty(R) + P_\infty(F) = 1$$

$$\Rightarrow P_\infty(R) = 2/3$$

6a) $P_0(S_2) = 1 - P_0(S_1) = 0.7$

b) $P(x_t | e_{1:t}) = \alpha P(e_t | x_t) \sum_{x_{t-1}} P(x_t | x_{t-1}) P(x_{t-1}, e_{1:t-1})$

$$P_1(S_1 | E_2) = \alpha P(E_2 | S_1) \cdot \Big(P(S_{1,\,t=1} | S_{1,\,t=0}) \cdot P(S_{1,\,t=0} | E_{2,\,t=0}) + P(S_{1,\,t=1} | S_{2,\,t=0}) \cdot P(S_{2,\,t=0} | E_{2,\,t=0}) \Big)$$

$$= 0,6 \cdot (0,4 \cdot 0,3 + 0,2 \cdot 0,7)$$

$$= 0,6 (0,12 + 0,14) = 0,6 \cdot 0,26 = 0,156$$

~~b) $P_1(S_1 | E_2) = P(S_1)$~~

b) $P_1(S_1) = P(E_2 | S_1) \cdot P'_1(S_1) / P(E_2)$

$$P'_1(S_1) = P(S_{1,\,t=1} | S_{1,\,t=0}) \cdot P_0(S_1) + P(S_{1,\,t=1} | S_{2,\,t=0}) \cdot P_0(S_2)$$

$$= 0,4 \cdot 0,3 + 0,2 \cdot 0,7 = 0,12 + 0,14 = \underline{0,26}$$

$$P_1(S_1) = 0,6 \cdot 0,26 / (0,26 \cdot 0,6 + 0,74 \cdot 0,4)$$

$$= \frac{39}{113} \approx 0,345$$

$$P_1(S_2) = \frac{113 - 39}{113} = 74/113 \approx 0,655$$

$39/113$

6c) $P_2'(S_1) = 0{,}4 \cdot \cancel{\ldots} + 0{,}2 \cdot \frac{74}{113} = \frac{152}{565}$

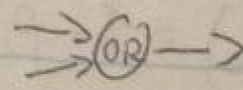
$$P_2(S_1) = P(E_3 \mid S_1) \cdot P_2'(S_1) \,/\, P(E_2)$$

$$= 0{,}2 \cdot \frac{152}{565} \Big/ \left(0{,}2 \cdot \frac{152}{565} + 0{,}4\left(1 - \frac{152}{565}\right)\right)$$

$$= \frac{76}{489} \approx 0{,}155$$

$$P_2(S_2) = \frac{413}{489} \approx 0{,}845$$

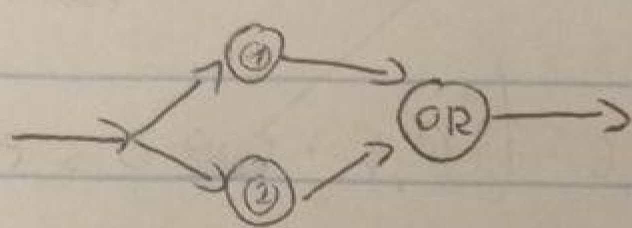
7) a) $w_1 = w_2 = 1$, $\theta = 1/2$



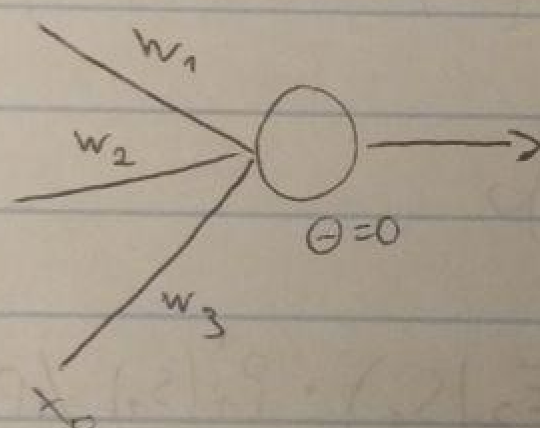
b) i) $w = 1$, $\theta = -10$

ii) $w = -1$, $\theta = -10$

c)



d)

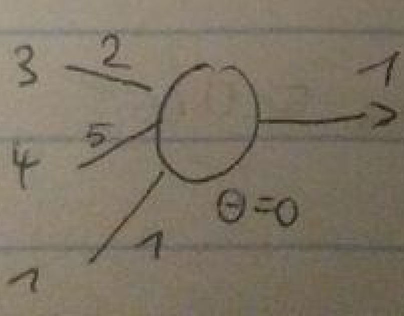


$$\vec{W}_0 := \vec{0}$$

$$\vec{w}_1 = \vec{w}_0 + c_0 \vec{x}_1 = \vec{x}_1 = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix}$$

$$\Rightarrow w_{1_1} = 2,\ w_{1_2} = 5,\ w_{1_3} = 1$$

e)

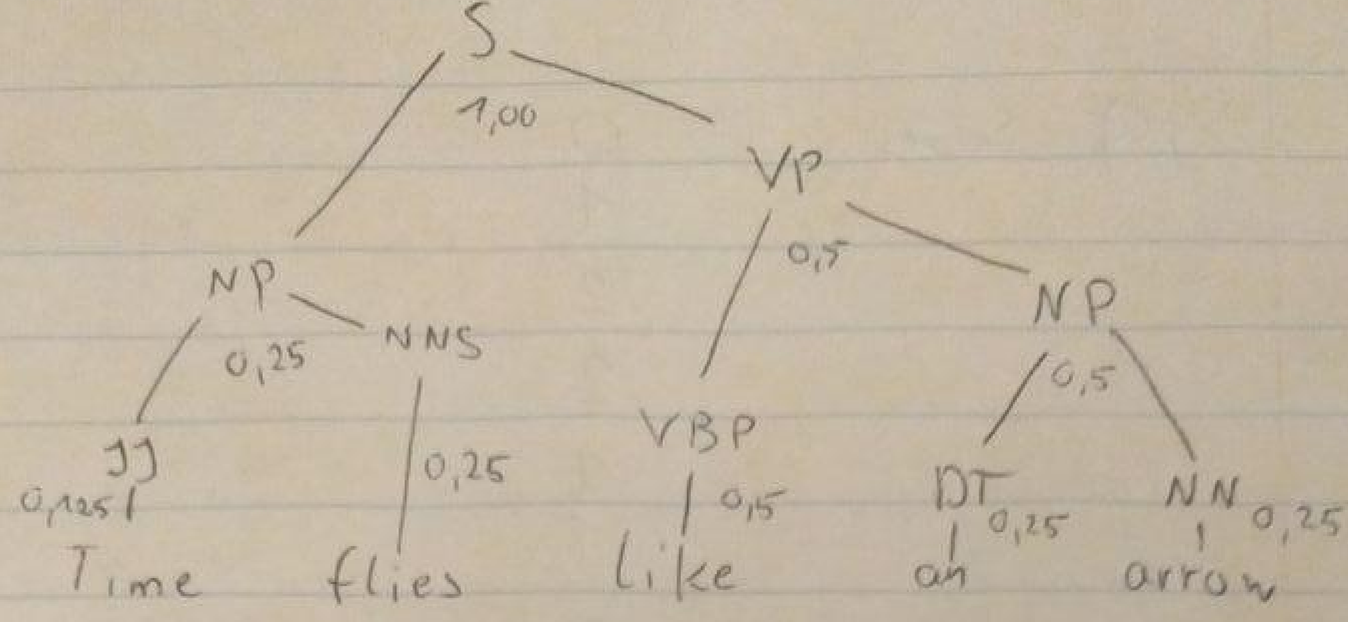


$$\Rightarrow \vec{w}_2 = \vec{w}_1 - \vec{x}_2 = \begin{pmatrix} 2 \\ 5 \\ 1 \end{pmatrix} - \begin{pmatrix} 3 \\ 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

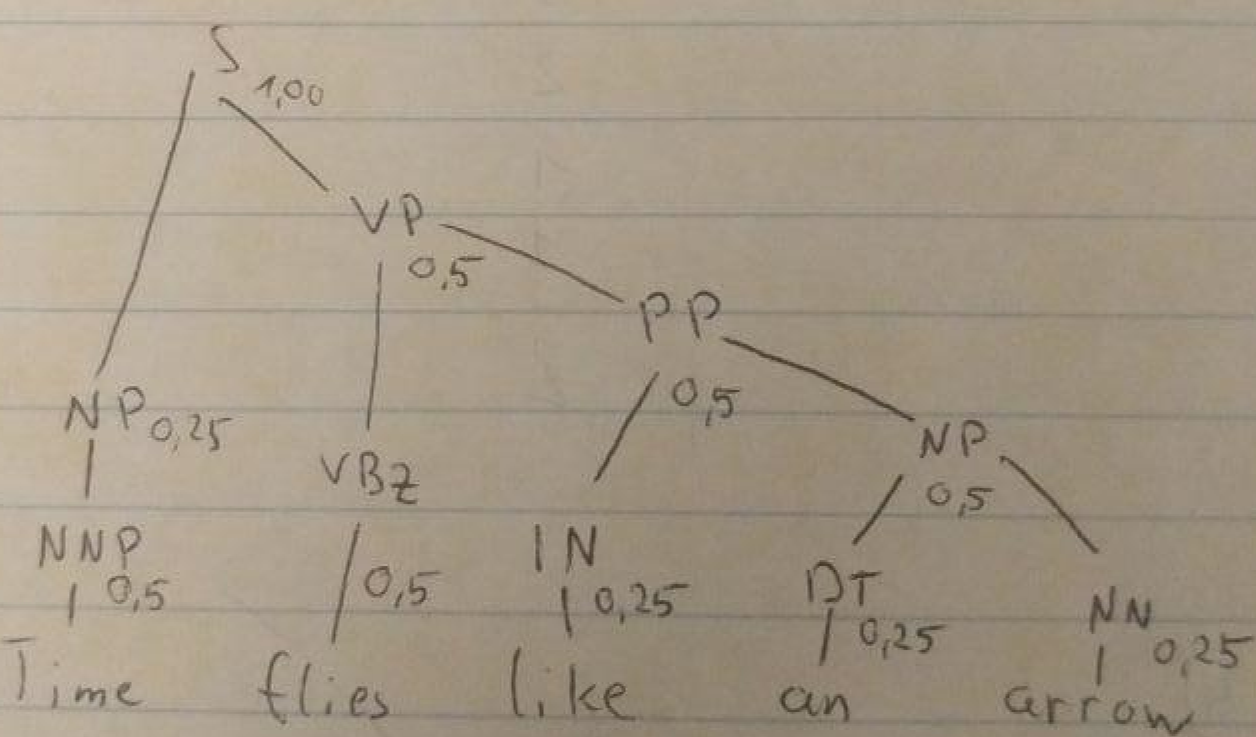
8) a) Scores:

0
1
2
2
1
1
2
1
2
1
3
2
1
1
2

b)



c)



d) $2+3+2+1+1+1+2+2 = 14$

e) $2+1+1+1+1+2+1+2+2 = 13$

f) parse 2 is more probable ($p_2 = 2^{-13}$, $P_1 = 2^{-14}$)

9)

a) A. Computer often think mechanical without ~~creativity~~ creative

B. Computer creative mechanical, think

C. mechanical Computer move part

b) {Computer, creative, mechanical, move, often, part, think, without}

c) A $(1,1,1,0,1,0,1,1)$

B. $(1,1,1,0,0,0,1,0)$

C. $(1,0,1,1,0,1,0,0)$

d) $\cos(\measuredangle(A,B)) = \frac{A \cdot B}{\|A\|_2 \cdot \|B\|_2} = \frac{4}{\sqrt{6}\ \sqrt{4}} = \frac{2}{\sqrt{6}}$

$\cos(\measuredangle(A,C)) = \frac{2}{\sqrt{6}\ \sqrt{4}} = \frac{1}{\sqrt{6}} < \cos(\measuredangle(A,B))$

e) A is more similar to B than to C.

10)

a) 1st law: A robot may never hurt or injure a human through action or inaction

2nd law: A robot must obey orders by humans, unless it conflicts with law 1.

3rd law: A robot must always protect itself unless it conflicts law 1 or 2.

b) The robots might refuse buying unhealthy food, because doing so would hurt the human, conflicting law 1.

c) I think the robots are applying the laws correctly, because doing otherwise might harm the human. Essentially this comes down to the question, if a robot might hurt a human, if the ~~s~~ same human wants it. According to Asimov's laws, this is not allowed.

d) When driving, the car might arrive in a situation, where it has to decide between hurting the passenger (crash into a wall) or hurting pedestrians (running them over). Since both options violate law 1, the car has to avoid such (and similar) situations by not driving at all.

e) Asimov's laws provide a basic framework for AI
moral, but following them as rigorously as in the
given story is not a desirable result. They might
be used as guidelines for AI design, but not as
laws written down in stories.