$\mathbb{E}[V] = \sum_{x \in Dom(X)} P(x)V(x)$

$\mathbb{E}[V(Decision)] = \sum_{outcome} P(outcome|decision)\, V(outcome)$

sum out what we can't control

Agents $\xrightarrow{\text{perform}}$ Actions $\xrightarrow{\text{results in}}$ Outcome

must. Even doing nothing is a choice

Usually not given. Only a chance of happening.

Lottery

Distribution over outcomes

$[p_1 : o_1, \dots, p_k : o_k]$

sum to 1

They **prefer** better outcomes

They are **rational**: do action w/ best outcome

Preference

For any 2 outcomes o_1, o_2, either

Weak Preference	$o_1 \succeq o_2$	"at least as desireable"
Strong Preference	$o_1 \succ o_2$	"more desireable"
Indifferent	$o_1 \sim o_2$	"as desireable"

properties:

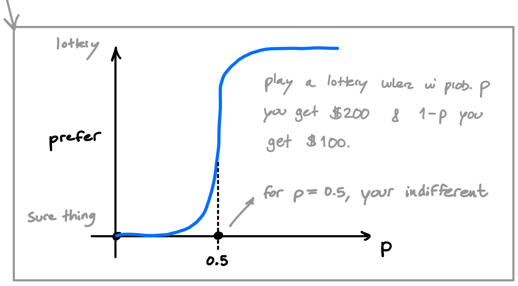
Complete $\quad o_1 \succeq o_2$ or $o_2 \succeq o_1$

Transitive $\quad o_1 \succeq o_2 \wedge o_2 \succeq o_3 \Rightarrow o_1 \succeq o_3$

prefers larger chance of better outcome

Monotonic $\quad o_1 \succeq o_2 \ \& \ p > q \Rightarrow [p : o_1, 1-p : o_2] \succeq [q : o_1, 1-q : o_2]$

Continuity $\quad o_1 \succ o_2 \wedge o_2 \succ o_3 \Rightarrow o_2 \sim \exists p, [1-p : o_1, p : o_3]$



Given these properties we can define...

Utility Function:

Map preference to numerical values

Utility : Outcome $\rightarrow$ [0, 1]

doesn't need to be in this range. But can always normalize

such that

① $o_1 \succeq o_2 \Leftrightarrow U(o_1) \geq U(o_2)$

Linear with Prob.s

② $U([p_1 : o_1, \dots, p_k : o_k]) = p_1 U(o_1) + \dots + p_k U(o_k)$

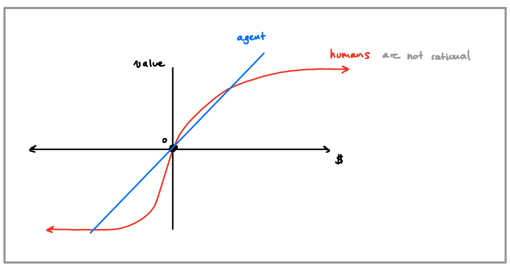
Rational Agents

Act to maximize expected utility.

a_1 results in $\qquad a_2$ results in

$U([p_1 : o_1, \dots, p_k : o_k]) \geq U([q_1 : o_1, \dots, q_k : o_k])$

agent chooses a_1



Decision Theory

How to **tradeoff** preferences & probability of outcomes, between actions.

Agent $\longrightarrow$ action?

depends on

ability | beliefs | preferences (& tradeoffs)

Std. Bayes Net

Decision Network

specified by

decision Variables | Utility

Decision Variables

Agent $\xrightarrow{\text{chooses value}}$ $D = d_i$ $\xrightarrow{\text{leads to}}$ ω outcome

expected utility: $\mathcal{E}(u|D=d_i) = \sum P(\omega|D=d_i)\, u(\omega)$

optimal Decision: $\arg\max_{d_i} \mathcal{E}(u|D=d_i)$

Decision Networks

For sequential decision problems

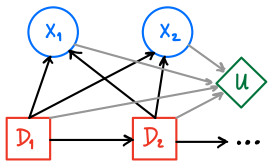
? $\longrightarrow$ X $\quad$ P(X|?)

? $\longrightarrow$ $\boxed{D_i}$ $\quad$ Info Available when decision D is made (when agent has to act)

? $\longrightarrow$ u $\quad$ Utility depends on

$$\mathcal{E}(u|D) = \sum_{X_1\dots X_n} P(X_1,\dots,X_n|D)\,U(X_1,\dots,X_n,D)$$

$$\prod_{j=1}^{n} P(X_j|D)$$



Optimal Decision:

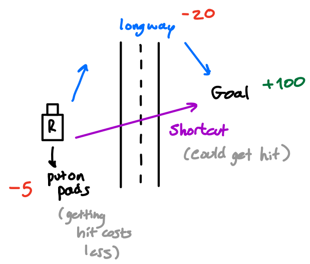
① Factor for each ___ and for ___

② Multiply, sum out : $\sum$ ___ ___ $\rightarrow$ factor for $\mathcal{E}(u|D)$

③ Choose d_i with max value in that factor

Same as BN-VE but:
- ordering is fixed
- Utility factor
- Maximization step.

Delivery Robot



Decision Var's :

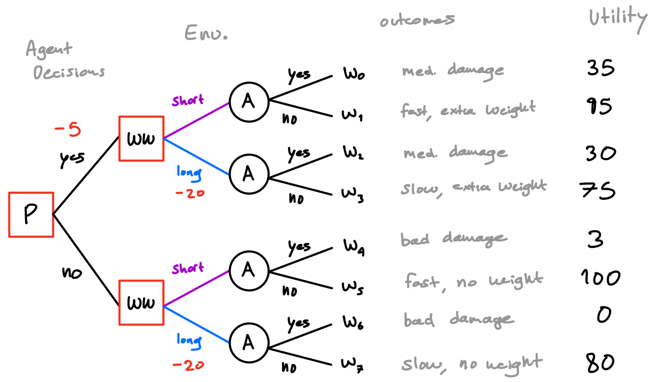
① Put on pads (P) : yes/no

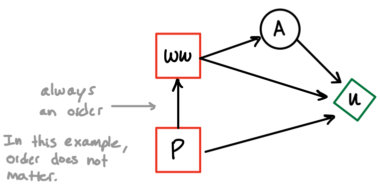
② Which way (ww) : long/short

Environment Var's :

① Accident (A) : yes/no

ww	A	P(A|ww)
long	yes	0.01
long	no	0.99
Short	yes	0.2
Short	no	0.8



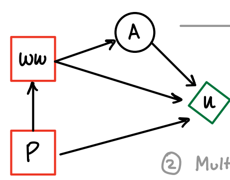


$$\mathcal{E}[u|P,ww] = \sum_{a\in A} P(A=a|P,ww)\,u(P,a,ww)$$

w.w	wp	$\mathcal{E}[u]$
s	T	83
l	T	74.6
s	F	80.6
l	F	79.2

As factors:

① Create Factors



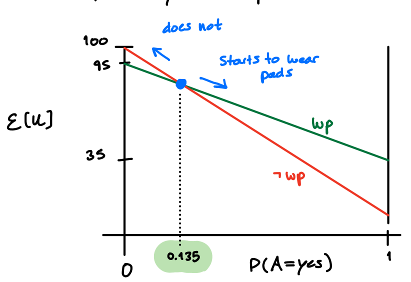
$f_1(A,ww) = P(A|ww)$

$f_2(A,ww,p) = P(u|A,ww,p)$

② Multiply & Sum Out

$$f_3(u,ww,p) = \sum_{A} f_1(A,ww)\,f_2(A,ww,p)$$
$$= \mathcal{E}(u|ww,p)$$

What probability of A, p would make the robot put on the pads?



$35p + (1-p)95 = 3p + (1-p)100$

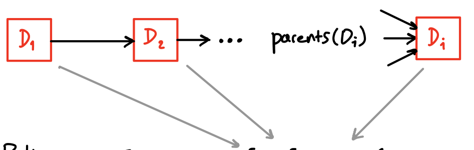
accident

no accident

$p = 0.135$

Sequential Decisions $D_1, \dots D_n$

Each D_i has an information set $parents(D_i)$ whose values will be know when its time to make D_i



Actions depend on previous observations

Policy Sequence $\delta_1, \delta_2, \dots, \delta_n$

dom[parents(D_i)] dom[D_i]

$$\delta_i(o) = d_i$$

expected Utility

$$\boxed{\mathcal{E}(u \mid \delta) = \sum_{\omega \models \delta} P(\omega)\, u(\omega)}$$

all worlds possible with this policy

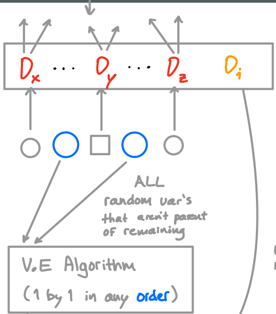
Algorithm: Find δ^* (optimal policy)

Init:

Factors: each CPT, Utility Function

remaining D_i's $\leftarrow$ $\boxed{D_1, D_2 \dots D_n}$ all of them

Repeat:



f_{k+1}

Pick D_i w/ no children. $\leftarrow$ i.e. In reverse order.

Eliminate by maximizing.

$f_{k+2} = \max_{D_i} f_{k+1}$ $\mathcal{E}(u \mid D_i)$

$\delta_i^* = \arg\max_{D_i} f_{k+1}$ (optimal for D_i) $\rightarrow$ $\delta_1, \delta_2, \dots, \delta_i$

Note:

Pool does not reset.

Add f_{k+1} to pool.

$\downarrow \emptyset$

End:

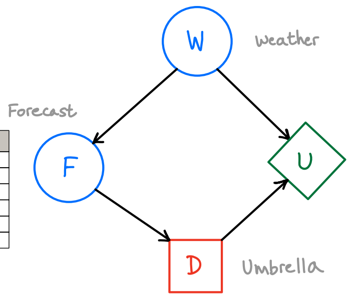
① Eliminate remaining variables $\rightarrow \dfrac{\mathcal{E}(u \mid \delta^*)}{P(\text{Evidence})}$

② Divide by P(Evidence) (if any)

Example – Umbrella

W	P(W)
clear	0.7
rain	0.3

$f_1(W)$



W	F	P(F|W)
clear	sun	0.7
clear	cloud	0.2
clear	rain	0.1
rain	sun	0.15
rain	cloud	0.25
rain	rain	0.6

$f_2(W,F)$

W	D	U
clear	yes	20
clear	no	100
rain	yes	70
rain	no	0

$f_0(W,D)$

$$\mathcal{E}(U|D) = \sum_{W,F} P(W,F)\,U(W,D) = \sum_{W,F} f_0(W,D)\, f_1(W)\, f_2(W,F)$$

---

① Eliminate W (VE)

$$f_3(F,D) = \sum_{W} f_0(W,D)\, f_1(W)\, f_2(W,F) =$$

F	D	f_3
sun	yes	12.95
sun	no	49
cloud	yes	8.05
cloud	no	14
rain	yes	14
rain	no	7

② Max out D

$$f_4(F) = \max_{D} f_3(F,D) =$$

F	D	f_4
sun	yes	12.95
sun	no	49
cloud	yes	8.05
cloud	no	14
rain	yes	14
rain	no	7

$$\therefore\ \delta_D(F) =$$

F	$\delta(F)$
sun	no
cloud	no
rain	yes

Sum-out F :

$$f_5(\,) = \sum_{F} f_4(F) = 49 + 14 + 14 = 77$$

(expected) value of optimal δ

This time, we observe F = cloud

policy is D = no. But $\mathcal{E}(D=\text{no}) \neq 14$

because this value considers F = cloud being unlikely

$$\mathcal{E}(U|D) = \sum_{W} P(W, F=\text{cloud})\,U(W,D)$$

$$f_2 \leftarrow f_2(W, F=\text{cloud}) =$$

W	F	f_2
clear	sun	0.7
clear	cloud	0.2
clear	rain	0.1
rain	sun	0.15
rain	cloud	0.25
rain	rain	0.6

=

W	f_2
clear	0.2
rain	0.25

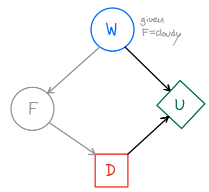
} unnormalized

$$f_1(W) \leftarrow P(W, F=\text{cloudy})\,P(W) =$$

W	f_1
clear	0.21
rain	0.075

(0.2) 0.7
(0.25) 0.3

} unnormalized



① Eliminate W (VE)

$$f_7(D) = \sum_{W} f_0(W,D)\, f_6(W) =$$

D	f_7
yes	8.05
no	14

② Max out D

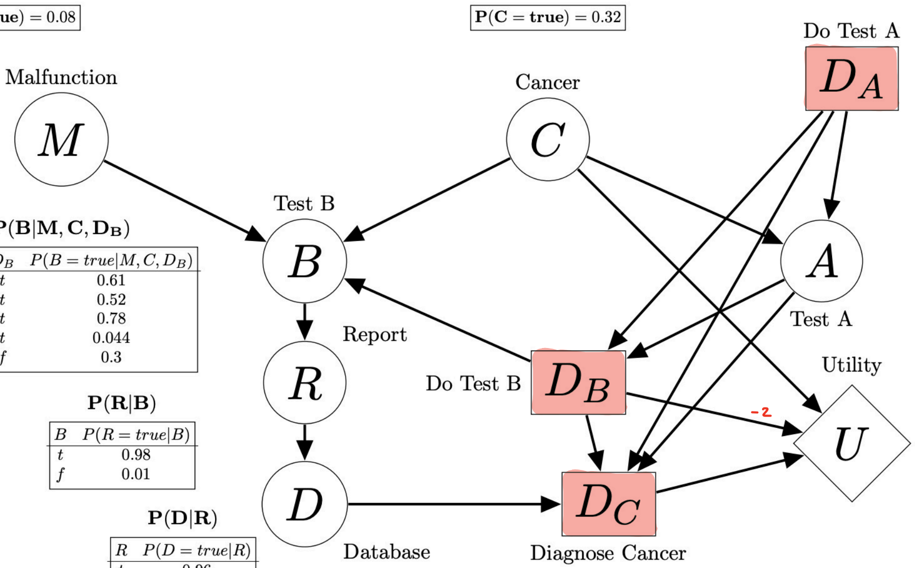
$$f_8(\,) = \max_{D} f_3(D) = 14 \qquad \therefore\ \delta_D = \text{no}$$

Divide by P(Evidence) to get expected value:

$$\sum_{W} P(W, F=\text{cloudy})\,P(W) = \underbrace{(0.2)0.7}_{W=\text{clear}} + \underbrace{(0.25)0.3}_{W=\text{rain}} = 0.215 \longrightarrow \frac{14}{0.215} = 65$$

$\mathbf{P(M = true) = 0.08}$

$\mathbf{P(C = true) = 0.32}$



$\mathbf{P(B|M, C, D_B)}$

M	C	D_B	$P(B = true\|M, C, D_B)$
t	t	t	0.61
t	f	t	0.52
f	t	t	0.78
f	f	t	0.044
$-$	$-$	f	0.3

$\mathbf{P(R|B)}$

B	$P(R = true\|B)$
t	0.98
f	0.01

$\mathbf{P(D|R)}$

R	$P(D = true\|R)$
t	0.96
f	0.001

$\mathbf{P(A|C, D_A)}$

C	D_A	$P(A = true\|C, D_A)$
t	t	0.7
f	t	0.35
$-$	f	0.462

$\mathbf{U(C, D_B, D_C)}$

Max reward: 20

C	D_B	D_C	$U(U, D_B, D_C)$
t	t	t	18
t	t	f	0
t	f	t	20
t	f	f	2
f	t	t	2
f	t	f	18
f	f	t	4
f	f	f	20

TP
FN -18
TN
FP -16

Init: Factors:

$f_0(C) =$

C	f_0
t	0.32
f	0.68

$f_1(M) =$

M	F_1
t	0.08
f	0.92

$f_2(B, M, C, DB) =$

B	M	C	D_B	f_3
t	t	t	t	0.61
t	t	t	f	0.39
t	f	t	t	0.52
t	f	t	f	0.48
f	t	t	t	0.78
f	t	t	f	0.22
f	f	t	t	0.044
f	f	t	f	0.956
–	–	f	t	0.3
–	–	f	f	0.7

$f_3(R, B) =$

B	R	f_3
t	t	0.98
t	f	0.02
f	t	0.01
f	f	0.99

$f_4(D, R) =$

R	D	f_4
t	t	0.96
t	f	0.04
f	t	0.001
f	f	0.999

$f_5(A, C, D_A) =$

C	DA	A	f_5
t	t	t	0.7
t	t	f	0.3
f	t	t	0.35
f	t	f	0.65
–	f	t	0.462
–	f	f	0.538

$f_6(C, D_B, D_C) =$

C	D_B	D_C	U
t	t	t	18
t	t	f	0
t	f	t	20
t	f	f	2
f	t	t	2
f	t	f	18
f	f	t	4
f	f	f	20

---

1. **Remaining:** D_A, D_B, D_C

No child in $\underline{\uparrow}$: M, C, B, R

No children: D_C

Variable Elim:

$$f_7(B, C, D_B) = \sum_M f_1(M)\, f_2(B, M, C, D_B)$$

$$f_8(R, C, D_B) = \sum_B f_3(R, B)\, f_7(B, C, D_B)$$

$$f_9(D, C, D_B) = \sum_R f_4(D, R)\, f_8(R, C, D_B)$$

$$f_{10}(A, D, D_A, D_B, D_C) = \sum_C f_0(C)\, f_5(A, C, D_A)\, f_6(C, D_B, D_C)\, f_9(D, C, D_B)$$

Pool

$f_0(C)$ $\cancel{f_2(B, M, C, DB)}$ $f_4(D, R)$ $\cancel{f_1(M)}$ $f_3(R, B)$ $f_5(A, C, D_A)$ $f_6(C, D_B, D_C)$

$f_0(C)$ $\cancel{f_3(R, B)}$ $f_4(D, R)$ $f_5(A, C, D_A)$ $f_6(C, D_B, D_C)$ $\cancel{f_7(B, C, D_B)}$

$f_0(C)$ $\cancel{f_4(D, R)}$ $f_5(A, C, D_A)$ $f_6(C, D_B, D_C)$ $\cancel{f_8(R, C, D_B)}$

$\cancel{f_0(C)}$ $\cancel{f_5(A, C, D_A)}$ $\cancel{f_6(C, D_B, D_C)}$ $\cancel{f_9(D, C, D_B)}$

Max Out:

$$f_{11}(A, D, D_A, D_B) = \max_{D_C} f_{10}(A, D, D_A, D_B, D_C)$$ expected Utility

$$\delta_C^* = \arg\max_{D_C} f_{10}(A, D, D_A, D_B, D_C)$$ policy

2. **Remaining:** D_A, D_B

No child in $\underline{\uparrow}$: M, C, B, R, D

No children: D_B

Variable Elim: relevent: D

$$f_{12}(A, D_A, D_B) = \sum_D f_{11}(A, D, D_A, D_B)$$

Pool

$\cancel{f_{11}(A, D, D_A, D_B)}$

Max Out:

$$f_{13}(A, D_A) = \max_{D_B} f_{12}(A, D_A, D_B)$$

$$\delta_B^* = \arg\max_{D_B} f_{10}(A, D, D_A, D_B, D_C)$$

3. **Remaining:** D_A

No child in $\underline{\uparrow}$: All

Variable Elim: relevent: A

$$f_{14}(D_A) = \sum_A f_{13}(A, D_A)$$

Pool

$\cancel{f_{13}(A, D_A)}$

Max Out:

$$f_{15}(\,) = \max_{D_A} f_{14}(D_A)$$

$$\delta_B^* = \arg\max_{D_A} f_{14}(D_A)$$

Final Policy

do test A
    if A= (+)
        do test B:
        diagnose based on D
    else:
        diagnose no cancer

Final Expected Value: 15.9