Planning under uncertainty

Thursday, November 14, 2019 2:19 PM

- Have to handle uncertainty:

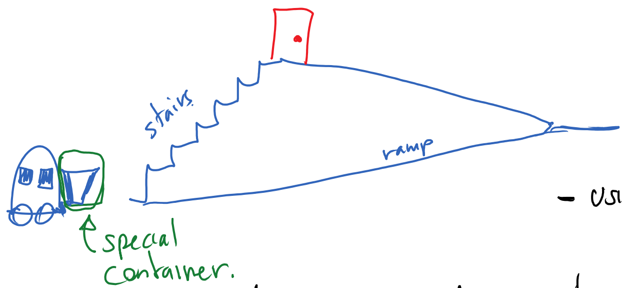
- Act regardless of certainty
  - preference over outcomse
  - Act to obtain prefered outcome.

- Our Agent can only know a probability distribution over outcomes.

- Modeling Preference:

Utility function u: world $\to \mathbb{R}$ $[0 \dots 100]$ $[-100 \dots 100]$

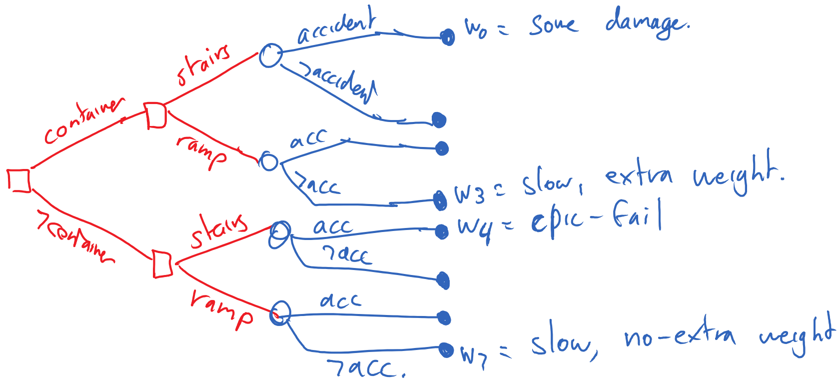
Example:



- ramp is slow but safe.
- stairs are quick but risky.
- using container adds weight.

2 decitions and one random outcome = accident.

decision tree:



$u(w) \Rightarrow [0 \dots 100]$

decicion D

def: Expected utility of a decision $D = d_i$

$$\mathcal{E}(u \mid D = d_i) = \sum_{w \in \Omega} P(w \mid D = d_i) * u(w)$$

weighted average by probability.

Optimal Decision

$$\mathcal{E}(u \mid D = d_{max}) = \max_{d \in Domain(D)} \mathcal{E}(u \mid D = d_i).$$

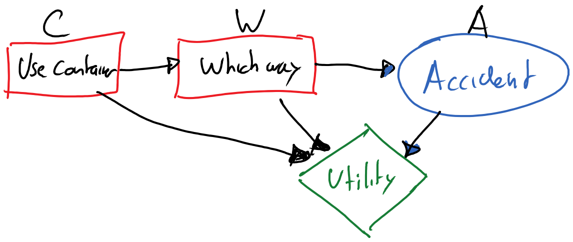
Difficult for Humans.- (not intuitive)

- Scratcher

Incorporate to Belief Networks: Decision Network:

A belief network with extra node types.

- Decision Nodes : rectangle Nodes that represent decisions of the agent
- Utility Node : Diamond Nodes that represent utilities, parents of an utility node will be the variables over which the utility depends.



decisions are given a total order

Tables required
- P(Accident | Which Way)
- u(Accident, Container, Which way)

$\Rightarrow$ d_{max}

no tables associated with decision nodes.

P(A | W)

W	A	
ramp	t	0.01
ramp	f	0.99
stairs	t	0.2
stairs	f	0.8

u()

W	C	A	u(w,c,A)
r	T	t	0
r	T	F	75
r	F	t	30
r	F	F	80
s	T	t	35
s	T	F	95
s	F	t	3
s	F	F	100

Optimal decision.

$\mathcal{E}(u \mid w, c)$

W	C	$\mathcal{E}(u \mid w,c)$
r	T	74.55 = 0.01·0 + 0.99·75
r	F	79.2 = 0.01·30 + 0.99·80
s	T	83.0 = 0.2·35 + 0.8·95
s	F	80.6 = 0.2·3 + 0.8·100

Decision $\rightarrow$ Effects $\rightarrow$ Utility

Now : Observe $\rightarrow$ Acts $\rightarrow$ Observe $\rightarrow$ Act.

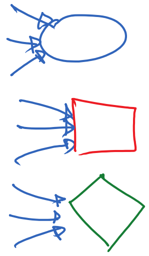
Extend belief networks to acomodate observe then act.

"Sequential Decision Problem"

Sequence of Decision variables $D_1, D_2, D_3 \ldots D_n$

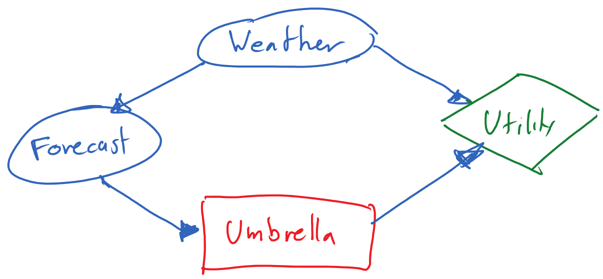
each D_i is going to have parents(D_i) "Information Set"

- Random Variables.



- Random Variables.
- Decision Variables.
- Utility Variables.

Example:



$W = \{rain, \neg rain\}$
$F = \{rain, cloudy, sunny\}$
$U = \{Yes, No\}$

- $P(W)$
- $P(F|W)$
- $u(W, U)$

W	P(W)
r	0.3
¬r	0.7

W	F	P(F|W)
¬r	s	0.7
¬r	c	0.2
¬r	r	0.1
r	s	0.15
r	c	0.25
r	r	0.60

W	U	Utility
¬r	Y	20
¬r	N	100
r	Y	70
r	N	0

Policies:
- always take umbrella
- never take umbrella.
- if forecast is rainy then take umbrella,
- if forecast is rainy or cloudy take the umbrella

Decision Function : A function that asigns to every decision D_i a value.

Policy : A sequence of decision functions π

- How to compare policies?

w.r.t Expected Utility $\mathcal{E}(u|\pi) = \sum_{w \text{ that satisfy } \pi} u(w) \cdot P(w)$

Optimal Policy π^* $\quad \mathcal{E}(u|\pi^*) > \mathcal{E}(u|\pi)$ for any other policy π

How to find such policy?
- compute $P(W, F)$ and check each utility

eleminate random variables that are not needed.

- create factors, eleminate random variables that are not needed.

$f_0(W)$ $\quad$ $f_4(W,F,U) = f_0(W)\cdot f_2(F,W)\cdot f_3(W,U)$

$f_1(F,W)$ $\quad$ $f_5(F,U) = \sum_W f_4(W,F,U)$

$f_3(W,U)$

- pick max rows from $f_5(F,U)$ → Policy π^* → Expected Utility π^*

$f_0(W)$

r	0.3
¬r	0.7

$f_1(F,W)$

r	s	0.7
r	c	0.2
r	r	0.1
¬r	s	0.15
¬r	c	0.25
¬r	r	0.6

$f_3(W,U)$

r	Y	70
r	N	0
¬r	Y	20
¬r	N	100

$f_4(W,F,U)$

r	s	Y	$= 0.3\cdot 0.7\cdot 70$
r	s	N	$= 0.3\cdot 0.7\cdot 0$
r	c	Y	$= 0.3\cdot 0.2\cdot 70$
r	c	N	
¬r	r	Y	$= 0.7\cdot 0.6\cdot 20$
¬r	r	N	$= 0.7\cdot 0.6\cdot 100$

$f_5(F,U)$

s	Y	12.95
s	N	~
c	Y	~
c	N	14.0
r	Y	~
r	N	7.0

select max

$f_6(F,U)$

S	N	49.00
C	N	14.00
r	Y	14.00

π^*

$E(U|\pi^*) = 77.00$

Example: