Decision Theory: Single Stage Decisions

Computer Science cpsc322, Lecture 33

(Textbook Chpt 9.2)

March, 30, 2009

Lecture Overview

- Intro
- One-Off Decision Example
- Utilities / Preferences and optimal Decision
- Single stage Decision Networks

Planning in Stochastic Environments

Environment

Problem		Deterministic	Stochastic
Static	Constraint Satisfaction	*Vars + Constraints*: Arc Consistency, Search, SLS	for CSP
	Query	*Logics*: Search (for Inference); CSP	*Belief Nets*: Var. Elimination; *Markov Chains and HMMs*
Sequential	Planning	*STRIPS*: Search (for complex planning); CSP	*Decision Nets*: Var. Elimination; *Markov Decision Processes*: Value Iteration

Representation

Reasoning Technique

Planning Under Uncertainty: Intro

- **Planning** how to select and organize a sequence of actions/decisions to achieve a given goal.
- **Deterministic Goal**: A possible world in which some proposition are true

- **Planning under Uncertainty**: how to select and organize a sequence of actions/decisions to "*maximize the probability*" of achieving a given goal
- **Goal under Uncertainty**: we'll move from all-or-nothing goals to a richer notion: rating how *happy* the agent is in different possible worlds.

“Single” Action vs. Sequence of Actions

Set of primitive decisions that can be treated as a single macro decision to be made *before acting*

- Agents makes observations
- Decides on an action
- Carries out the action

Sequential Decisions

Lecture Overview

- Intro
- One-Off Decision Example
- Utilities / Preferences and Optimal Decision
- Single stage Decision Networks

Recap: One-off decision example

Delivery Robot Example

not

- Robot needs to reach a certain room

A true false

- Going through stairs may cause an accident.
- It can go the short way through long stairs, or the long way through short stairs (that reduces the chance of an accident but takes more time)

Which way: short / long

$P(A \mid short) > P(A \mid long)$

- The Robot can choose to wear pads to protect itself or not (to protect itself in case of an accident) but pads slow it down

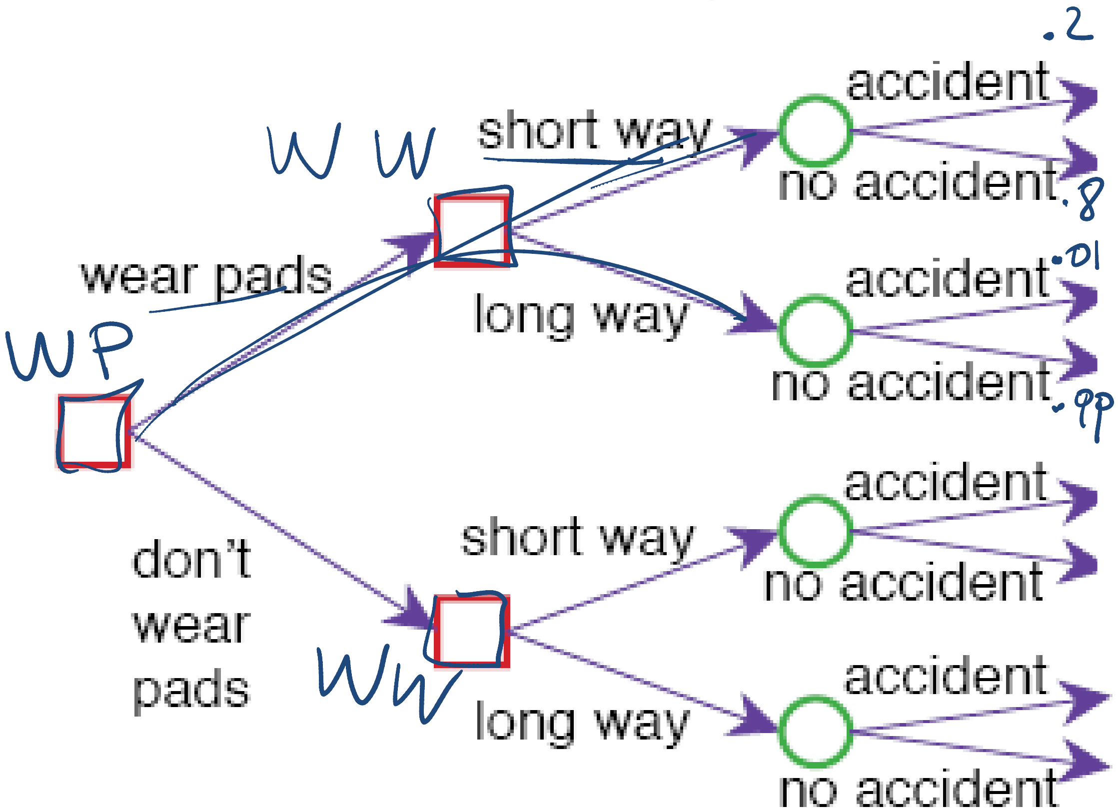
Wear Pads: true / false

If A=true

- If there is an accident the Robot does not get to the room

Decision Tree for Delivery Robot

- This scenario can be represented as the following decision tree

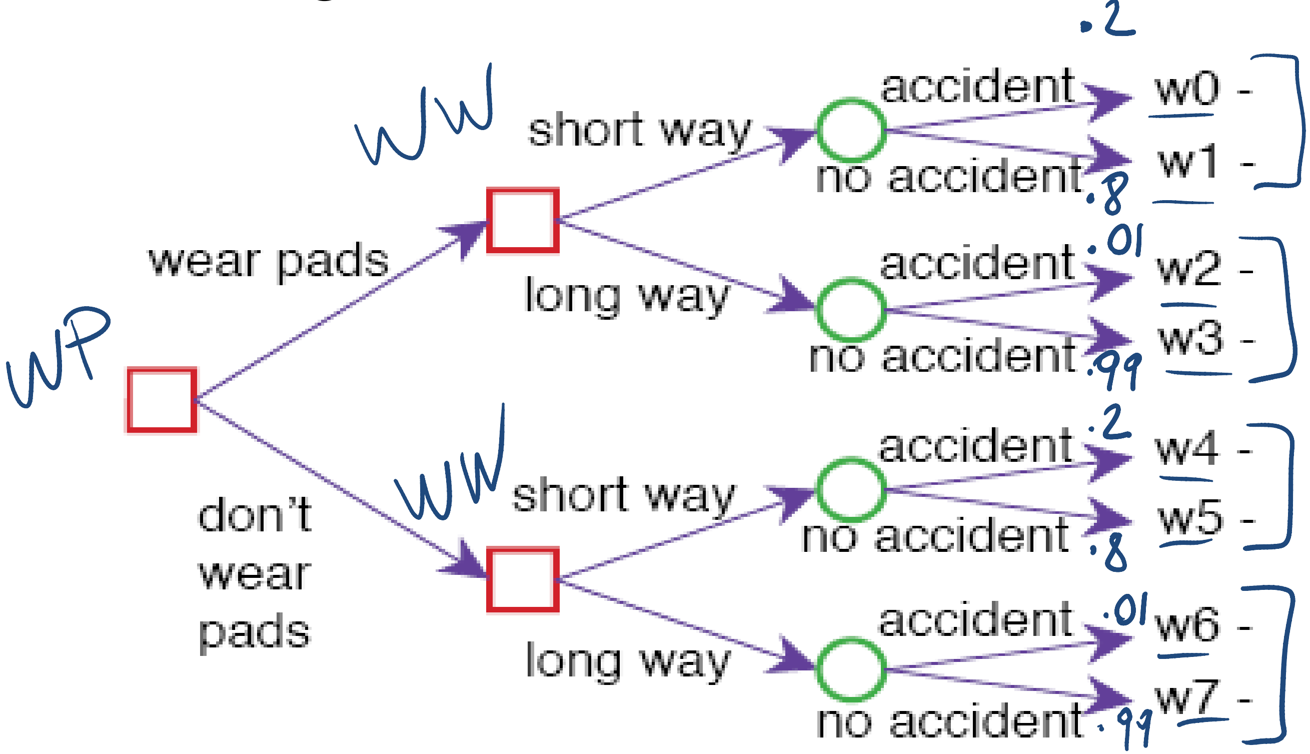


Which way	Accident	
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8

- The agent has a set of decisions to make (a macro-action it can perform)
- Decisions can influence random variables
- Decisions have probability distributions over outcomes

Decision Variables: Some general Considerations

- A possible world specifies a value for each random variable and each decision variable.
- For each assignment of values to all decision variables, the probabilities of the worlds satisfying that assignment sum to 1.



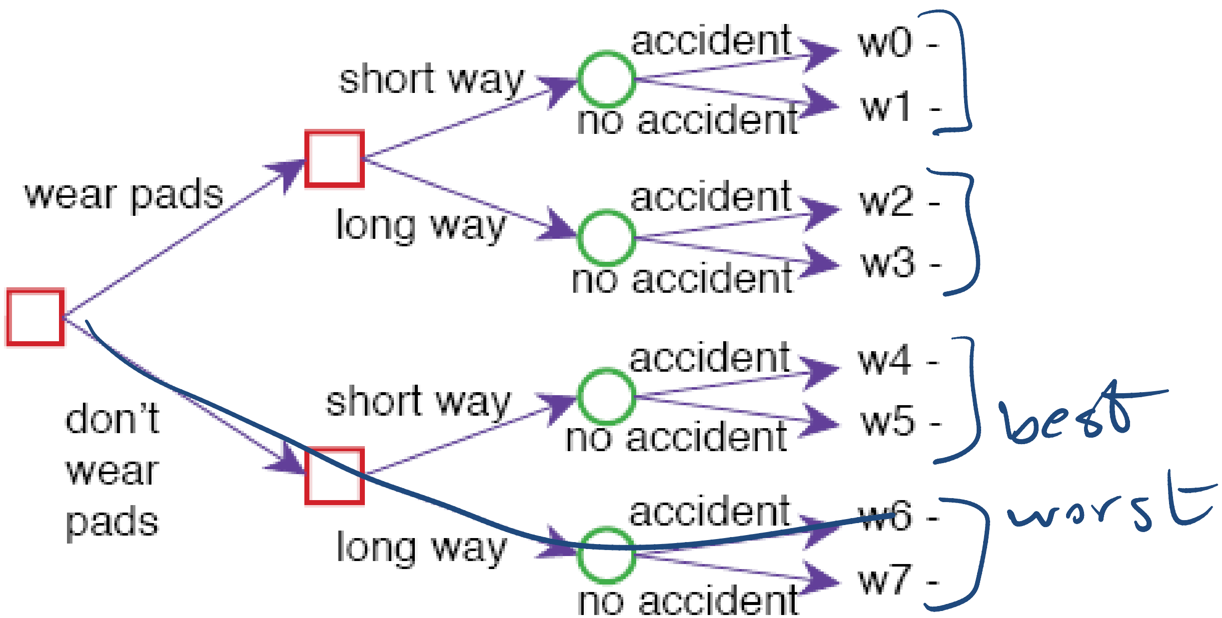
Lecture Overview

- Intro
- One-Off Decision Problems
- Utilities / Preferences and Optimal Decision
- Single stage Decision Networks

What are the optimal decisions for our Robot?

It all depends on how happy the agent is in different situations.

For sure getting to the room is better than not getting there….. but we need to consider other factors..



Utility / Preferences

Utility: a measure of desirability of possible worlds to an agent

- Let U be a real-valued function such that $U(w)$ represents an agent's degree of preference for world w . [0, 100]

This would be a reasonable utility function for our Robot

Which way	Accident	Wear Pads	Utility	World
short	true	true	35	w0, moderate damage
short	false	true	95	w1, reaches room, quick, extra weight
long	true	true	30	w2, moderate damage, low energy
long	false	true	75	w3, reaches room, slow, extra weight
short	true	false	3	w4, severe damage
short	false	false	100	w5, reaches room, quick
long	false	false	0	w6, severe damage, low energy
long	true	false	80	w7, reaches room, slow

Utility: Simple Goals

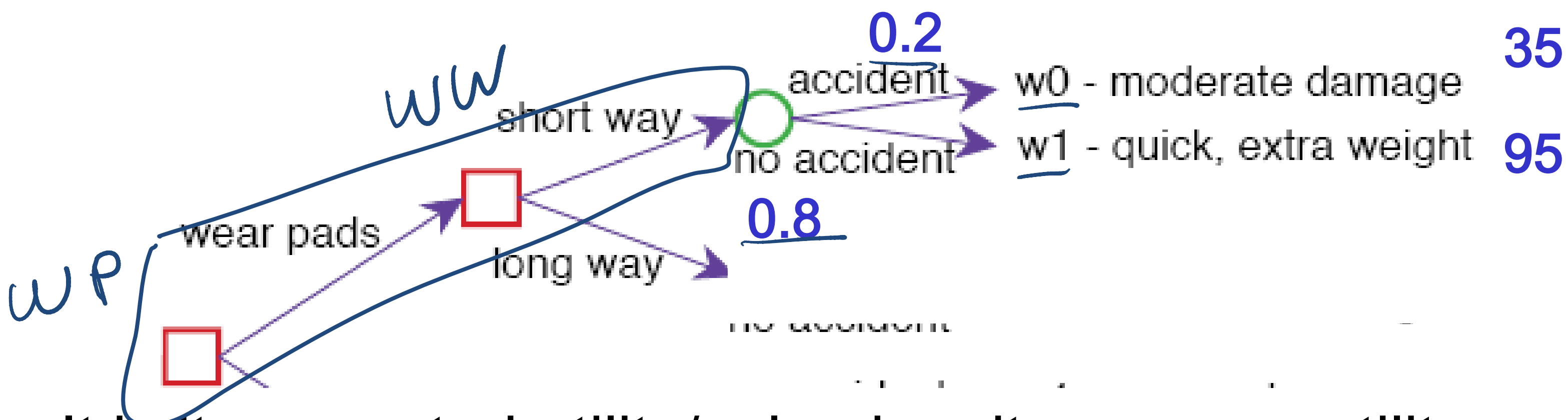
- Can simple (boolean) goals still be specified?

reading the room

Which way	Accident	Wear Pads	Utility
long	true	true	0
long	true	false	0
long	false	true	100
long	false	false	100
short	true	true	0
short	true	false	0
short	false	true	100
short	false	false	100

Optimal decisions: How to combine Utility with Probability

What is the utility of achieving a certain probability distribution over possible worlds?



- It is its expected utility/value i.e., its average utility, weighting possible worlds by their probability.

$EU(WP = t, WW = short) =$

$= 0.2 * 35 + .8 * 95$

Optimal decision in one-off decisions

WP

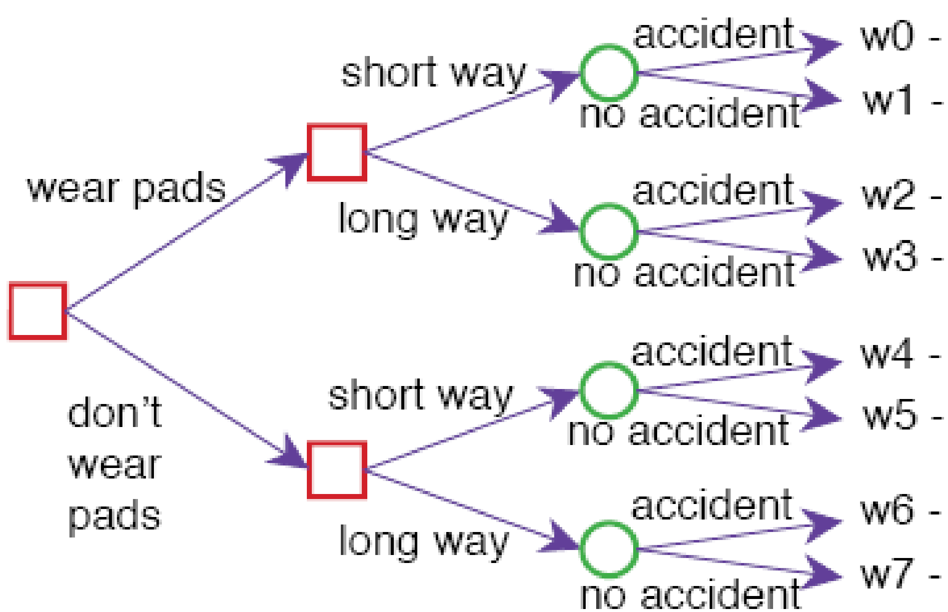
- Given a set of n decision variables var_i (e.g., Wear Pads, Which Way), the agent can choose:

W

$\boxed{D = d_i}$ for any $d_i \in \text{dom}(var_1) \times .. \times \text{dom}(var_n)$.

d_i

Wear Pads	Which way
true	short
true	long
false	short
false	long

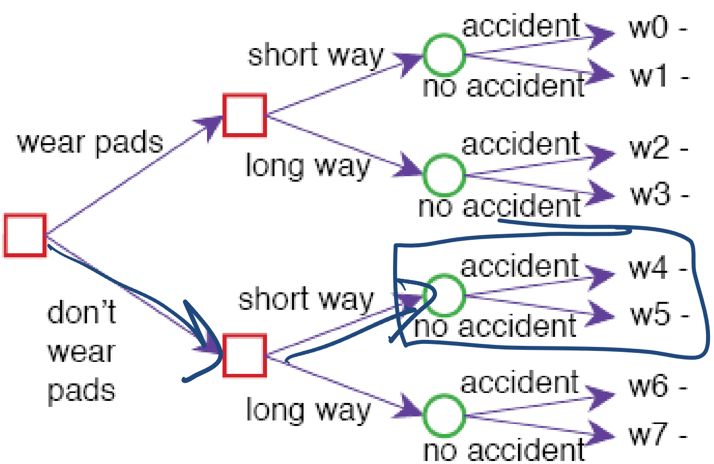


Optimal decision: Maximize Expected Utility

- The expected utility of decision $D = d_i$ is

$$\mathbb{E}(U \mid D = d_i) = \sum_{w \models D = d_i} P(w \mid D = d_i)\, U(w)$$

e.g., $\mathbb{E}(U \mid D = \{WP = \text{false}, WW = \text{short}\}) =$



$P(w_4 \mid D) * U(w_4) +$
$+ P(w_5 \mid D) * U(w_5)$

- An optimal decision is the decision $D = d_{max}$ whose expected utility is maximal:

$$d_{max} = \underset{d_i \in dom(D)}{\arg\max}\, \mathbb{E}(U \mid D = d_i)$$

max

Wear Pads	Which way	EU
true	short	—
true	long	—
false	short	—
false	long	—

Lecture Overview

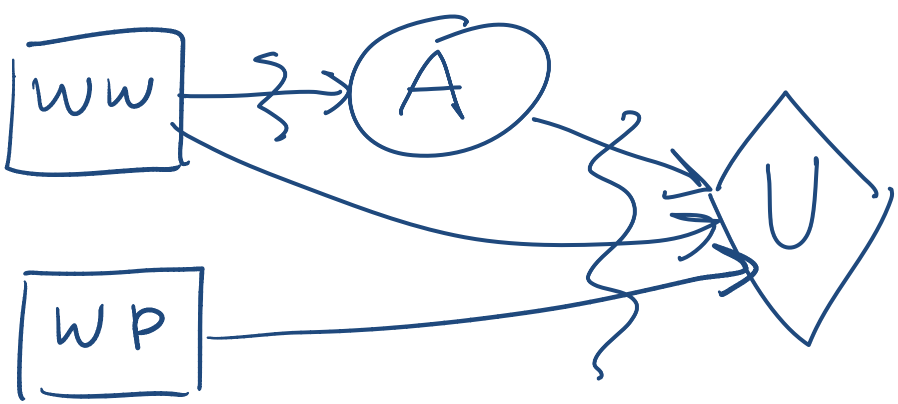
- Intro
- One-Off Decision Problems
- Utilities / Preferences and Optimal Decison
- Single stage Decision Networks

Single-stage decision networks

Extend belief networks with:

- **Decision nodes**, that the agent chooses the value for. Drawn as rectangle.
- **Utility node**, the parents are the variables on which the utility depends. Drawn as a diamond.
- Shows explicitly which decision nodes affect random variables

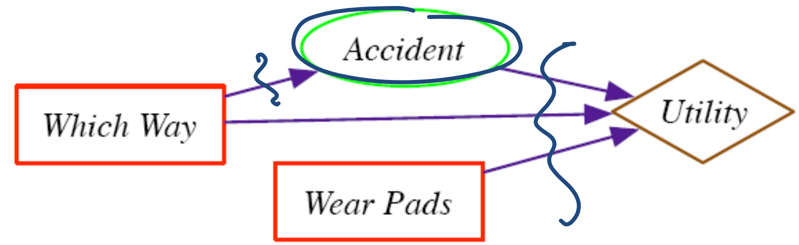
Which way	Accident	
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8



Which way	Accident	Wear Pads	Utility
long	true	true	30
long	true	false	0
long	false	true	75
long	false	false	80
short	true	true	35
short	true	false	3
short	false	true	95
short	false	false	100

Finding the optimal decision: We can use VE

Suppose the random variables are $X_1, \ldots, X_n$, the decision variables are the set D, and utility depends on $pU \subseteq \{X_1, \ldots, X_n\} \cup D$



$$\mathbb{E}(U \mid D) = \sum_{X_1,\ldots,X_n} P(X_1,\ldots,X_n \mid D)\, U(pU)$$

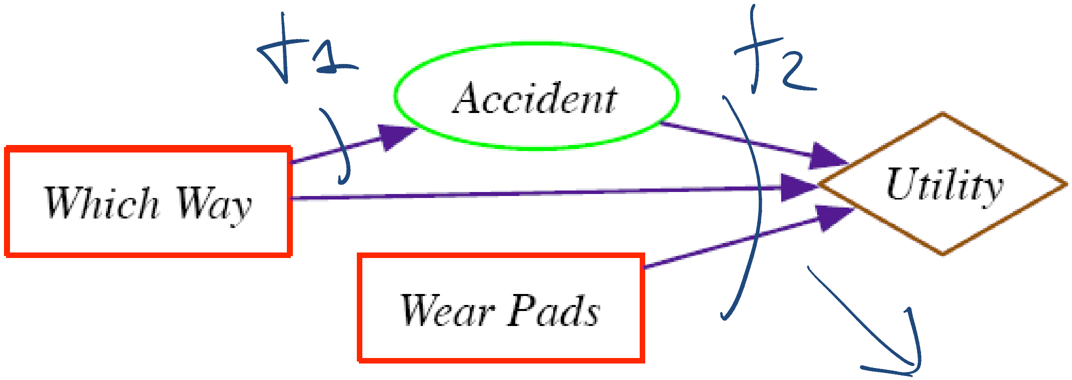
$$= \sum_{x_i} \prod P(x_i \mid pX_i)\, U(pU)$$

include decision vars

To find the optimal decision we can use VE:

1. Create a factor for each conditional probability and for the utility
2. Multiply factors and sum out all of the random variables (This creates a factor on D that gives the expected utility for each D)
3. Choose the D with the maximum value in the factor.

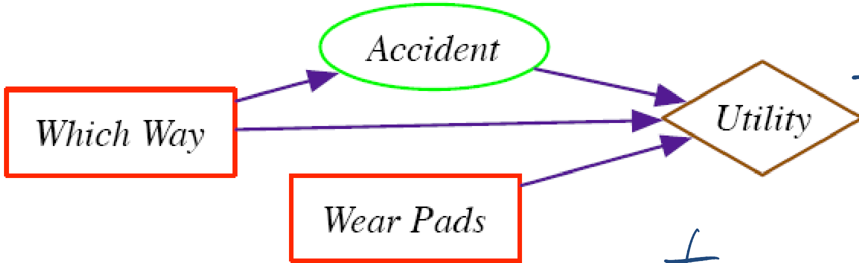
Example Initial Factors (Step1)



Which way	Accident	Probability
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8

Which way	Accident	Wear Pads	Utility
long	true	true	30
long	true	false	0
long	false	true	75
long	false	false	80
short	true	true	35
short	true	false	3
short	false	true	95
short	false	false	100

Example: Multiply Factors (Step 2a)



$$\sum_{A} f_1(WW, A) \times f_2(A, WW, WP)$$

$f_1 \times f_2$

f_2

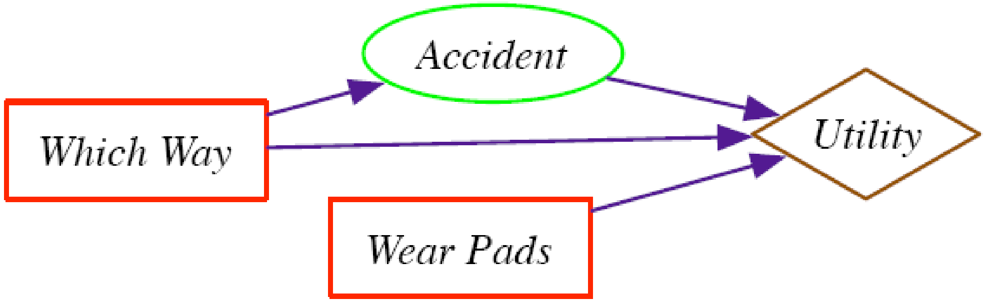
Which way	Accident	Probability
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8

f_2

Which way	Accident	Wear Pads	Utility
long	true	true	30
long	true	false	0
long	false	true	75
long	false	false	80
short	true	true	35
short	true	false	3
short	false	true	95
short	false	false	100

Which way	Accident	Wear Pads	Utility
long	true	true	30 * .01 ....
long	true	false	0 * .01
long	false	true	75 * .99
long	false	false	80 * .99
short	true	true	35
short	true	false	3
short	false	true	95
short	false	false	100

Example: Sum out vars and choose max (Steps 2b-3)



$$\sum_A f'(A, WW, WP)$$

A

Sum out accident: f''

Which way	Accident	Wear Pads	Utility
long	true	true	0.01*30
long	true	false	0.01*0
long	false	true	0.99*75
long	false	false	0.99*80
short	true	true	0.2*35
short	true	false	0.2*3
short	false	true	0.8*95
short	false	false	0.8*100

WW WP

Which way	Wear Pads	Expected Utility
long	true	0.01*30+0.99*75=74.55
long	false	0.01*0+0.99*80=79.2
short	true	0.2*35+0.8*95=83
short	false	0.2*3+0.8*100=80.6

Thus the optimal policy is to take the short way and wear pads, with an expected utility of 83.

Learning Goals for today's class

You can:

- Compare and contrast stochastic **single-stage (one-off)** decisions vs. **multistage** decisions
- Define a **Utility Function** on possible worlds
- Define and compute **optimal one-off decision** (max expected utility)
- Represent one-off decisions as **single stage decision networks** and compute optimal decisions by Variable Elimination

Next Class (textbook sec. 9.3)

one-off

Set of primitive decisions that can be treated as a single macro decision to be made *before acting*

Sequential Decisions

- Agents makes observations
- Decides on an action
- Carries out the action