Decision Theory: Single Stage Decisions

CPSC 322 Lecture 32

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**

Lecture Overview

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

R&R systems we'll cover in this course

		Environment	
Problem		Deterministic	Stochastic
Static	Constraint Satisfaction	*Variables + Constraints* Search Arc Consistency Local Search	
	Query	*Logics* Search	*Bayesian (Belief) Networks* Variable Elimination
Sequential	Planning	*STRIPS* Search	*Decision Networks* Variable Elimination

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 propositions 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

“Single” or **“One-off”**: Set of one or more primitive decisions that can be treated as a single macro decision to be made ***before*** acting, with no prior observations

“Sequential”: Set of one or more decisions, each of which depends on observations

- Agents makes observations
- Decides on an action
- Carries out the action
- *Repeats with future decisions*

Lecture Overview

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

One-off decision example

Delivery Robot Example

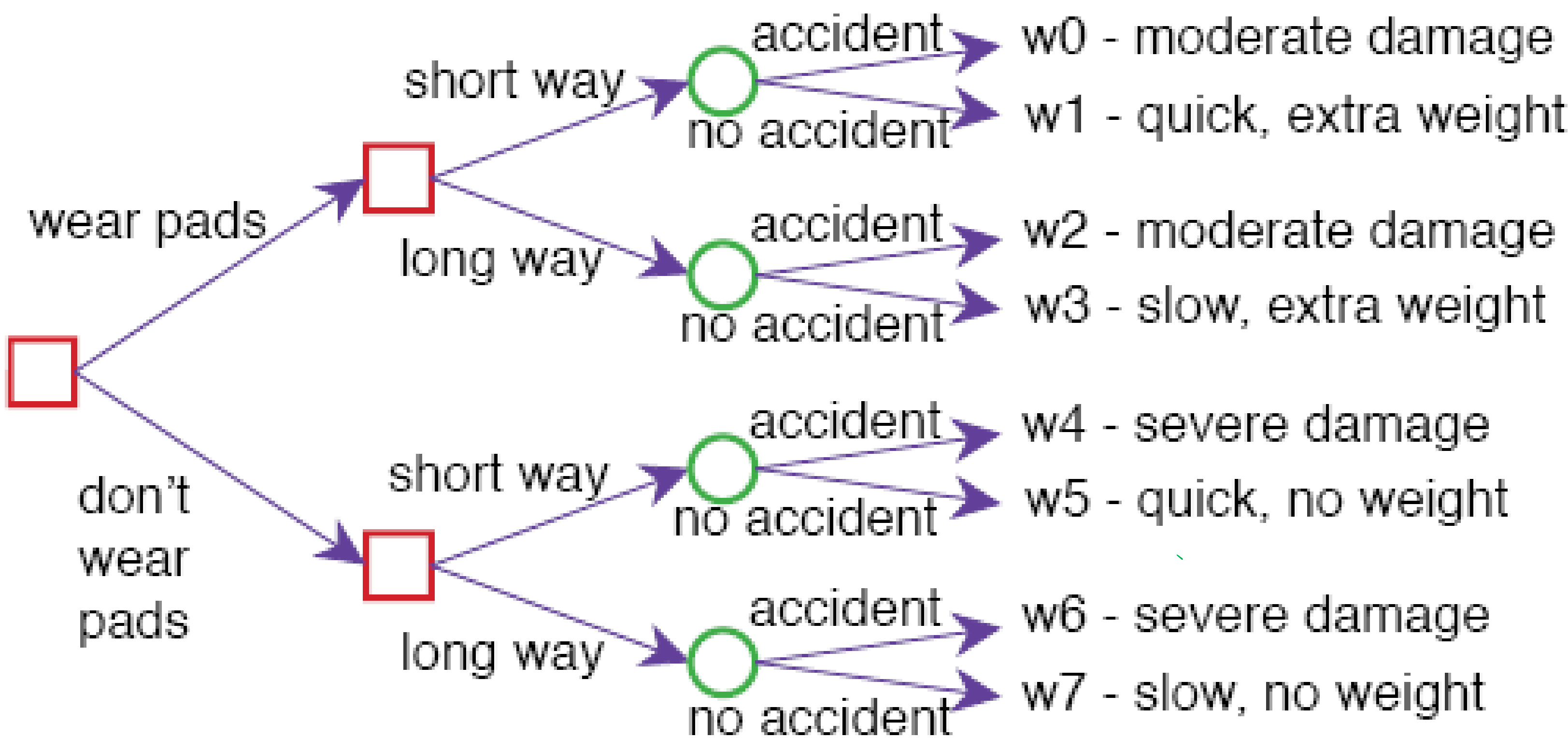
- Robot needs to reach a certain room
- Going through stairs may cause an accident.
- It can go the short way down the stairs, or the long way down a long ramp (that reduces the chance of an accident but takes more time)

- The Robot can choose to wear pads (to reduce damage in case of accident) or not, but pads slow it down

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

Decision Tree for Delivery Robot

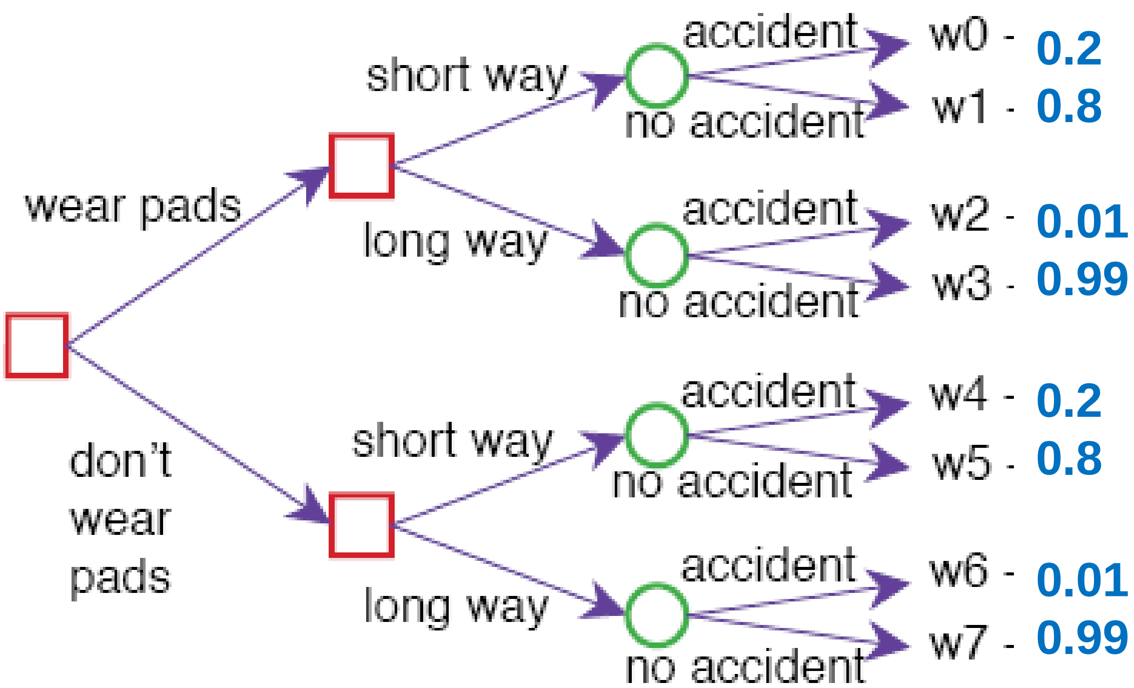
- This scenario can be represented as the following decision tree



- 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: 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.



Which way	Accident	P(A|WW)
long	true	0.01
long	false	0.99
short	true	0.2
short	false	0.8

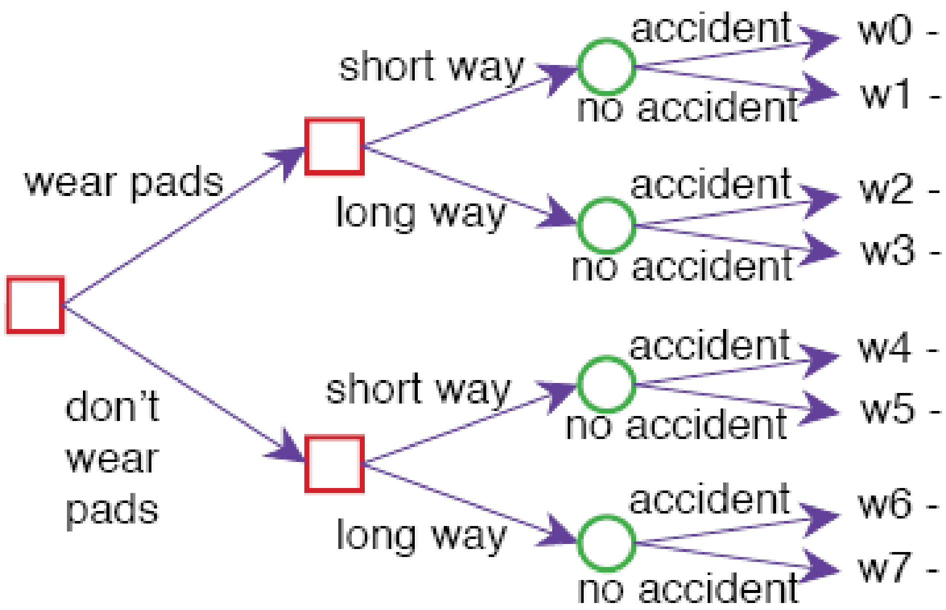
Lecture Overview

- 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* .

Would this be a reasonable utility function for our Robot, who wants to reach the room?

Which way	Accident	Wear Pads	Utility	World
short	true	true	18	w0, moderate damage
short	false	true	95	w1, reaches room, quick, extra weight
long	true	true	11	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	true	false	0	w6, severe damage, low energy
long	false	false	80	w7, reaches room, slow

A. Yes

B. No

C . I only care if this is on the exam

Utility: Simple Goals

- How can the simple (boolean) goal “reach the room” be specified?

A.

Which way	Accident	Wear Pads	Utility
long	true	true	0
long	true	false	0
long	false	true	0
long	false	false	0
short	true	true	0
short	true	false	0
short	false	true	100
short	false	false	90

B.

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

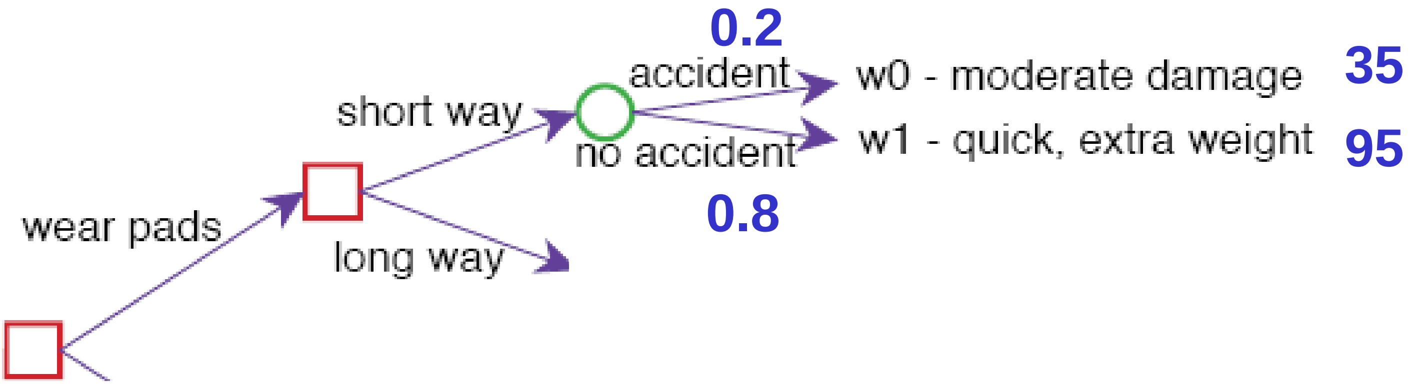
C .

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

D. ***Not possible***

Optimal decisions: Combining Utility & Probability

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



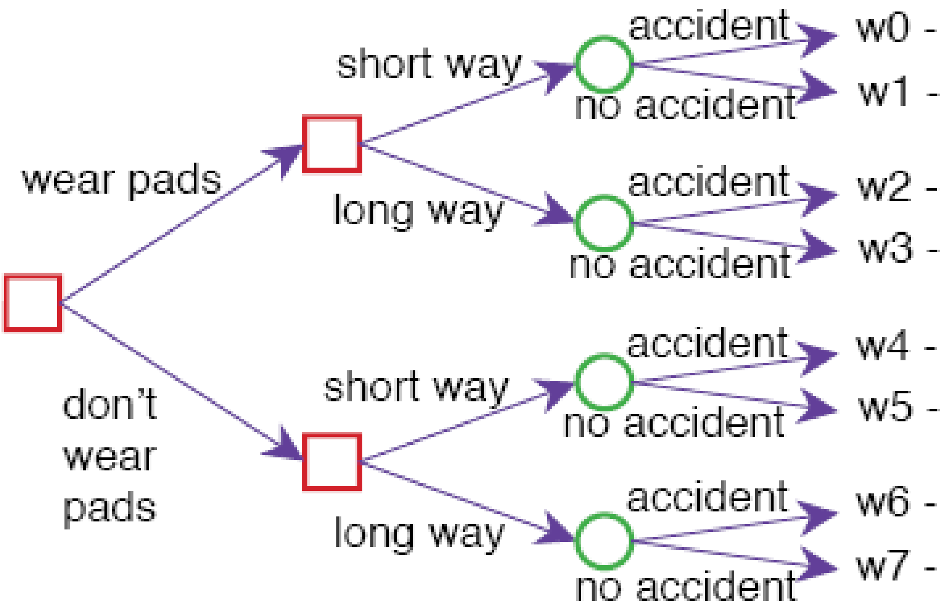
- It is its **expected utility/value** i.e., its “average” utility, weighting possible worlds by their probability.

Optimal decision in one-off decisions

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

  $D = d_i$ for any $d_i \in$ dom(var_1) x .. x dom(var_n) .

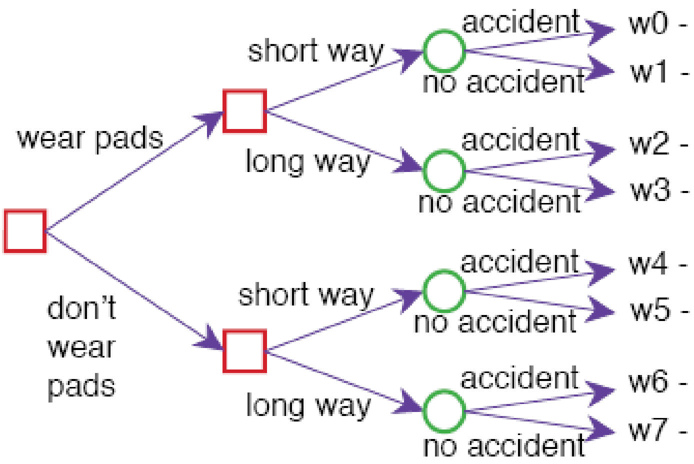
Wear Pads	Which way
true	short
true	long
false	short
false	long



Optimal decision: Maximize Expected Utility

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

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



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

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

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

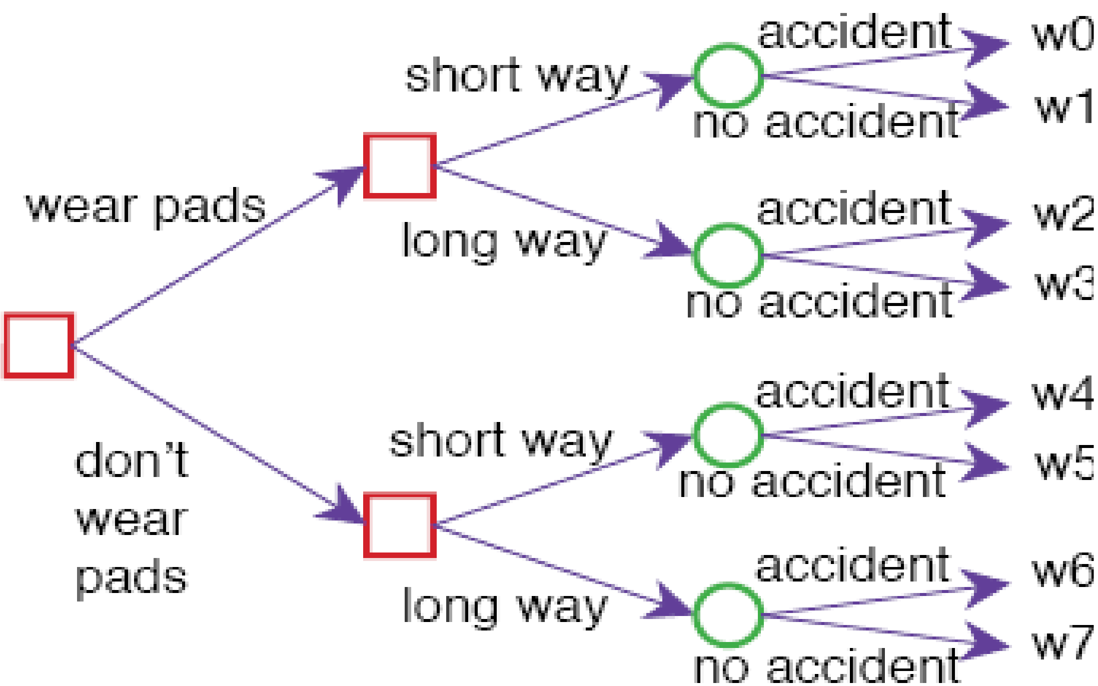
Expected utility of a decision

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

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

- What is the expected utility of Wearpads=true, WhichWay=short ?

0.2 * 35 + 0.8 * 95 = 83



Conditional probability	Utility	E[U|D]
0.2	35	83
0.8	95	
0.01	30	74.55
0.99	75	
0.2	3	80.6
0.8	100	
0.01	0	79.2
0.99	80	

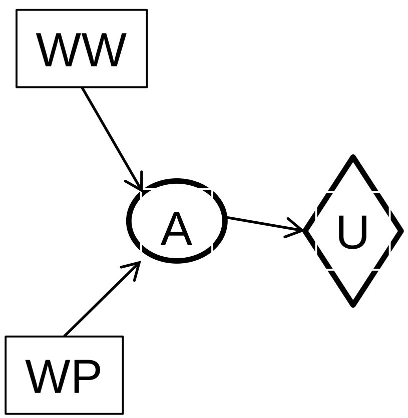
Lecture Overview

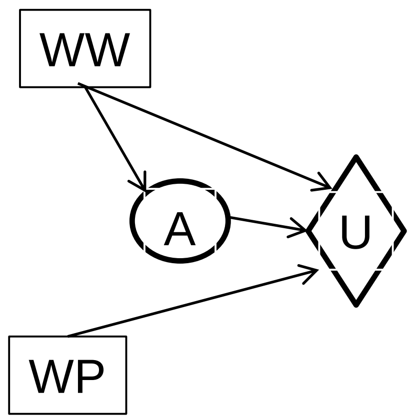
- Intro
- One-Off Decision Problems
- Utilities / Preferences and Optimal Decison
- Single stage Decision Networks

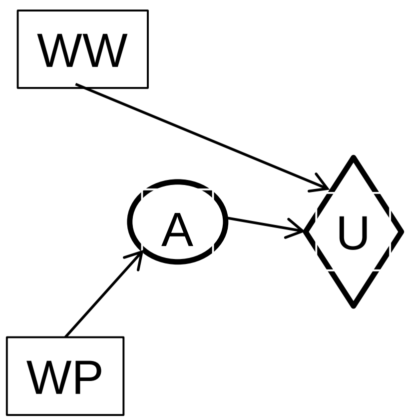
Single-stage decision networks

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







A.

B.

C .

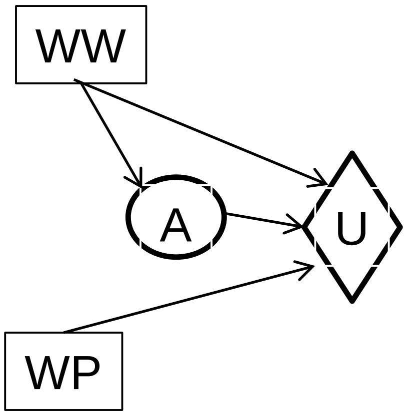
D . None of these

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

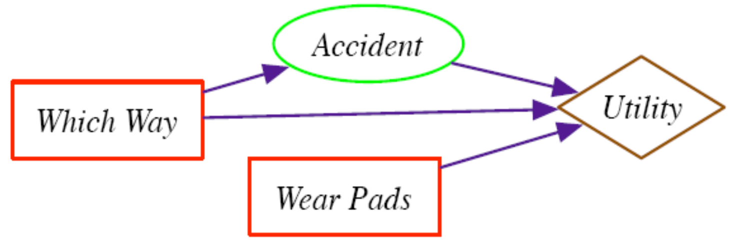


Which way	Accident	Wear Pads	Utility
long	true	true	30
long	true	false	0
long	false	true	75
long	false	false	80
short	true	true	35
short	true	false	3
short	false	true	95
short	false	false	100

Finding the optimal decision: Use VE

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

parents U



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

$$= \sum \prod P(x_i \mid pX_i)\, U(pU)$$

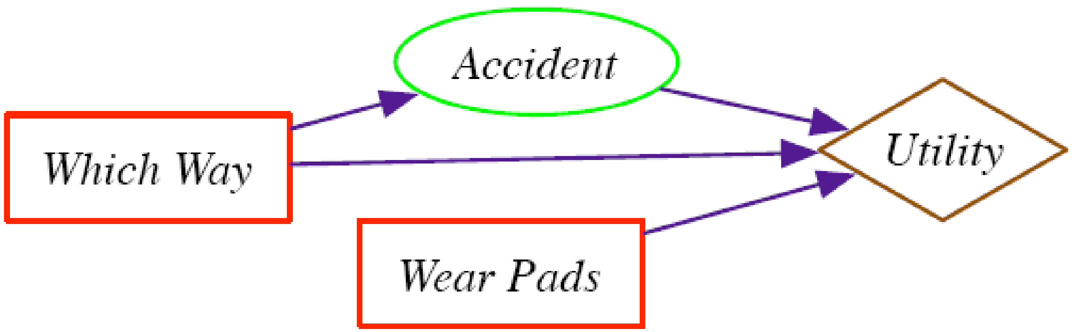
decision vars

also includes

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 that gives the expected utility for each decision)
3. Choose the decision with the maximum value in the factor.

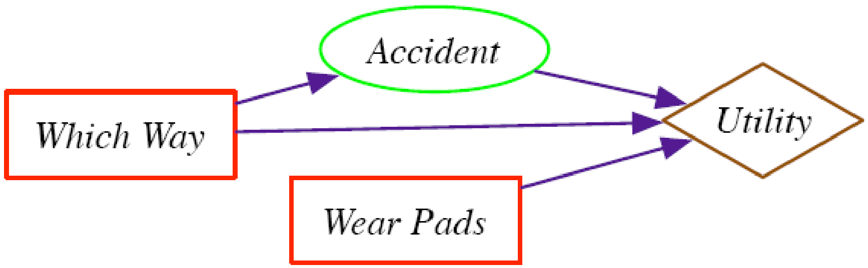
Example Initial Factors (Step1)



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)



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

f1

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

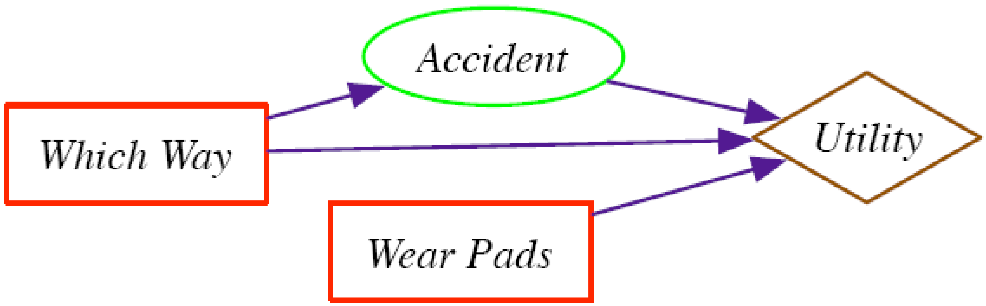
f2 →

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

f3

Which way	Accident	Wear Pads	Utility
long	true	true	30 * 0.01
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: Sum out vars and choose max (Steps 2b-3)



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

Sum out accident:

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

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.

Next Class (textbook sec. 9.3)

Sequential Decisions

- Agent makes observations
- Decides on an action
- Carries out the action
- *Repeats as appropriate for the situation*