

COMP111: Artificial Intelligence

Section 9b. Reasoning under Uncertainty 2b

Frank Wolter

Content

- ▶ (Conditional) Independence and Belief Networks
- ▶ Expected Value
- ▶ Expected Value and Decision Making

Independence to the Rescue

Random variables F and G are **independent** if

$$\mathbf{P}(F, G) = \mathbf{P}(F) \times \mathbf{P}(G),$$

that is, for all values r and s

$$P(F = r, G = s) = P(F = r) \times P(G = s)$$

As one's dental problems do not influence the weather, the pairs of random variables

- ▶ *Toothache, Weather,*
- ▶ *Catch, Weather*
- ▶ *Cavity, Weather*

are each independent.

Example: Weather and Dental Problems

The full joint probability distribution

$$\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity}, \textit{Weather})$$

has 32 entries. It contains four tables for the variables *Toothache*, *Catch*, *Cavity*. One for each kind of weather from

sunny, rain, cloudy, snow

Thus, we have

- ▶ 8 probabilities for (*Weather* = *sunny*):

$$P(\textit{Weather} = \textit{sunny}, \textit{Toothache} = r_1, \textit{Catch} = r_2, \textit{Cavity} = r_3)$$

- ▶ 8 probabilities for (*Weather* = *rain*):

$$P(\textit{Weather} = \textit{rain}, \textit{Toothache} = r_1, \textit{Catch} = r_2, \textit{Cavity} = r_3)$$

- ▶ and so on for (*Weather* = *cloudy*) and (*Weather* = *snow*).

What is the relationship between these editions?

Example: Weather and Dental Problems

Clearly we can make the independence assumption that for any combination of values of the random variables *Toothache*, *Catch*, *Cavity*, the the probabilities for *Weather*. For example,

$$\begin{aligned} &P(\textit{Weather} = \textit{sunny} | \textit{Toothache} = r_1, \textit{Catch} = r_2, \textit{Cavity} = r_3) \\ &= P(\textit{Weather} = \textit{sunny}) \end{aligned}$$

for all $r_1, r_2, r_3 \in \{0, 1\}$.

Thus, equivalently,

$$\begin{aligned} &P(\textit{Weather} = \textit{sunny}, \textit{Toothache} = r_1, \textit{Catch} = r_2, \textit{Cavity} = r_3) \\ &= P(\textit{Weather} = \textit{sunny})P(\textit{Toothache} = r_1, \textit{Catch} = r_2, \textit{Cavity} = r_3) \end{aligned}$$

for all $r_1, r_2, r_3 \in \{0, 1\}$.

The same equations hold for *rain*, *cloudy*, and *snow*.

Example: Weather and Dental Problems

We have seen that the joint probability distribution

$$\mathbf{P}(\textit{Weather}, \textit{Toothache}, \textit{Catch}, \textit{Cavity})$$

can be written as:

$$\mathbf{P}(\textit{Weather}) \times \mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$$

The 32-element table for four variables can be constructed from one 4-element table and one 8-element table.

Independence Analysis

- ▶ Dentist domain: the variables *Toothache*, *Catch*, and *Cavity* are all dependent on each other.
- ▶ Student exam domain with variables *Grade*, *Answers*, *Background*, *Works_hard*:
It seems reasonable to assume that *Background* and *Works_hard* are independent. No other pair is independent.
- ▶ Fire alarm domain: *Fire*, *Alarm*, *Tampering*, *Smoke*, *Leaving*, and *Report*.
It seems reasonable to assume that *Tampering* and *Fire*, and *Tampering* and *Smoke* are independent. No other pair is independent.

We conclude that independence is rare. We will now turn to conditional independence.

Conditional Independence

Random variables G, F are **conditionally independent given** H_1, \dots, H_n if

$$\mathbf{P}(G, F \mid H_1, \dots, H_n) = \mathbf{P}(G \mid H_1, \dots, H_n) \times \mathbf{P}(F \mid H_1, \dots, H_n)$$

or, equivalently,

$$\mathbf{P}(G \mid F, H_1, \dots, H_n) = \mathbf{P}(G \mid H_1, \dots, H_n)$$

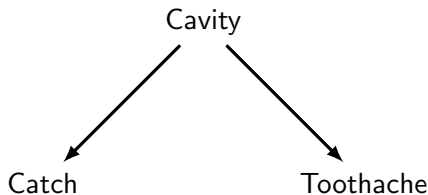
Example: Dentistry

In the dentist domain it seems reasonable to assert conditional independence of the variables *Toothache* and *Catch*, given *Cavity*:

$$\mathbf{P}(\textit{Toothache}, \textit{Catch} \mid \textit{Cavity}) = \mathbf{P}(\textit{Toothache} \mid \textit{Cavity})\mathbf{P}(\textit{Catch} \mid \textit{Cavity})$$

or, equivalently,

$$\mathbf{P}(\textit{Toothache} \mid \textit{Cavity}) = \mathbf{P}(\textit{Toothache} \mid \textit{Catch}, \textit{Cavity})$$



How does this help?

Example: Dentistry

Using conditional independence of *Catch* and *Toothache* given *Cavity* we can compute the joint probability distribution

$$\mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity})$$

using **only** the probability distributions:

$$\mathbf{P}(\textit{Toothache} \mid \textit{Cavity}), \quad \mathbf{P}(\textit{Catch} \mid \textit{Cavity}), \quad \mathbf{P}(\textit{Cavity})$$

The computation is as follows (using first multiplication rule and then conditional independence):

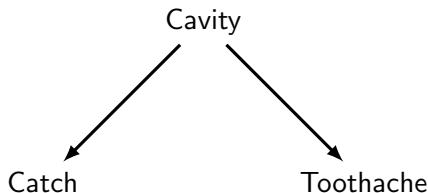
$$\begin{aligned} & \mathbf{P}(\textit{Toothache}, \textit{Catch}, \textit{Cavity}) \\ = & \mathbf{P}(\textit{Toothache}, \textit{Catch} \mid \textit{Cavity}) \times \mathbf{P}(\textit{Cavity}) \\ = & \mathbf{P}(\textit{Toothache} \mid \textit{Cavity}) \times \mathbf{P}(\textit{Catch} \mid \textit{Cavity}) \times \mathbf{P}(\textit{Cavity}) \end{aligned}$$

The number of probabilities needed is reduced to 5. Moreover, these probabilities can often be learned from data.

Towards Belief Networks

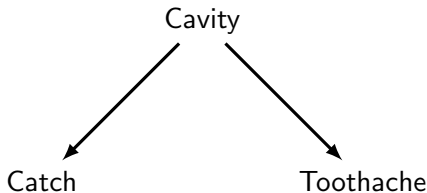
Conditional independence can be used to give concise representations of many domains.

A **belief network** (Bayesian network) is a graphical probabilistic model of a domain in which nodes represent random variables and arcs probabilistic dependence (often causality).



Towards Belief Networks

Informally, if there is an arc from a random variable F to another random variable G then G **depends on** F . F is called a **parent** of G . It is assumed that there are no cycles and that any random variable G is **conditionally independent** of any non-parent variable G' given the parents of G if G' cannot be reached by a sequence of arcs from G . For example:



The full joint probability distribution is then given as

$$\prod_{F \text{ in network}} \mathbf{P}(F \mid \text{parents}(F))$$

In the example $\mathbf{P}(\textit{Toothache} \mid \textit{Cavity}) \times \mathbf{P}(\textit{Catch} \mid \textit{Cavity}) \times \mathbf{P}(\textit{Cavity})$.

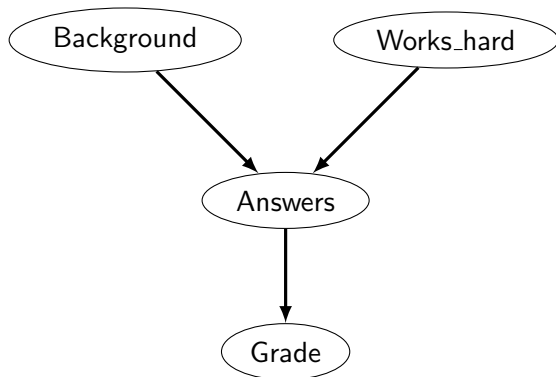
Student Exam Domain

Variables: *Grade*, *Answers*, *Background*, *Works_hard*.

Then it seems reasonable to assume that

- ▶ *Works_hard* and *Background* are independent;
- ▶ *Grade* and *Works_hard* are independent given *Answers* and *Grade* and *Background* are independent given *Answers*.

We represent this modeling of the domain using the Belief Network:



Fire Alarm Domain

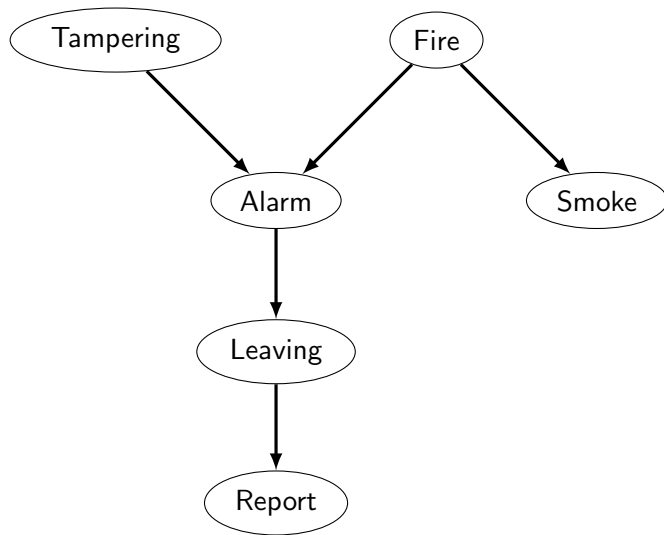
Recall the fire alarm system and reporting domain with random variables (all take value 0 or 1):

- ▶ *Fire*: there is fire;
- ▶ *Alarm*: the alarm goes off;
- ▶ *Tampering*: there is tampering with the alarm system;
- ▶ *Smoke*: there is smoke (no smoke detector used);
- ▶ *Leaving*: people leave the building;
- ▶ *Report*: it is reported that people leave the building (reporting not always correct).

Assume

- ▶ Fire is conditionally independent of Tampering;
- ▶ Alarm depends on Fire and Tampering;
- ▶ Smoke depends only on Fire and is conditionally independent of Tampering and Alarm given Fire;
- ▶ Leaving is conditionally independent of the other variables above given Alarm;
- ▶ Report only directly depends on Leaving.

Fire Alarm Domain



Joint Probability Distribution from $\mathbf{P}(F \mid \text{parents}(F))$

Given a Belief Network, we can always assume an ordering

$$F_1, \dots, F_n$$

of its random variables such that for all i, j :

$$F_i \rightarrow F_j \quad \text{implies} \quad i < j$$

In our example, we can order the random variables as follows:

Background, Works_hard, Answers, Grade

Tampering, Fire, Alarm, Smoke, Leaving, Report

Joint Probability Distribution from $\mathbf{P}(F \mid \text{parents}(F))$

According to the Chain Rule (proof on Exercise 8), given

F_1, \dots, F_n , we have for all r_1, \dots, r_n :

$$\begin{aligned} P(F_1 = r_1, \dots, F_n = r_n) &= P(F_1 = r_1) \times \\ &\quad P(F_2 = r_2 \mid F_1 = r_1) \times \\ &\quad P(F_3 = r_3 \mid F_1 = r_1, F_2 = r_2) \times \\ &\quad \dots \\ &\quad P(F_n = r_n \mid F_1 = r_1, \dots, F_{n-1} = r_{n-1}) \end{aligned}$$

Using bold \mathbf{P} notation this means:

$$\begin{aligned} \mathbf{P}(F_1, \dots, F_n) &= \mathbf{P}(F_1) \times \\ &\quad \mathbf{P}(F_2 \mid F_1) \times \\ &\quad \mathbf{P}(F_3 \mid F_1, F_2) \times \\ &\quad \dots \\ &\quad \mathbf{P}(F_n \mid F_1, \dots, F_{n-1}) \end{aligned}$$

Joint Probability Distribution from $\mathbf{P}(F \mid \text{parents}(F))$

$$\begin{aligned}\mathbf{P}(F_1, \dots, F_n) &= \mathbf{P}(F_1) \times \\ &\quad \mathbf{P}(F_2 \mid F_1) \times \\ &\quad \mathbf{P}(F_3 \mid F_1, F_2) \times \\ &\quad \dots \\ &\quad \mathbf{P}(F_n \mid F_1, \dots, F_{n-1})\end{aligned}$$

As $\text{parents}(F_i) \subseteq \{F_1, \dots, F_{i-1}\}$ conditional independence implies:

$$\mathbf{P}(F_i \mid F_1, \dots, F_{i-1}) = \mathbf{P}(F_i \mid \text{parents}(F_i))$$

Thus:

$$\begin{aligned}\mathbf{P}(F_1, \dots, F_n) &= \mathbf{P}(F_1) \times \\ &\quad \mathbf{P}(F_2 \mid \text{parents}(F_2)) \times \\ &\quad \mathbf{P}(F_3 \mid \text{parents}(F_3)) \times \\ &\quad \dots \\ &\quad \mathbf{P}(F_n \mid \text{parents}(F_n))\end{aligned}$$

Example: Student Exam Domain

The full joint probability distribution

$$\mathbf{P}(\textit{Background}, \textit{Works_hard}, \textit{Answers}, \textit{Grade})$$

can then be computed as

$$\begin{aligned} &\mathbf{P}(\textit{Background}) \times \\ &\mathbf{P}(\textit{Works_hard}) \times \\ &\mathbf{P}(\textit{Answers} \mid \textit{Background}, \textit{Works_hard}) \times \\ &\mathbf{P}(\textit{Grade} \mid \textit{Answers}) \end{aligned}$$

Example: Fire Alarm

The full joint probability distribution

$$\mathbf{P}(Tampering, Fire, Alarm, Smoke, Leaving, Report)$$

can then be computed as

$$\begin{aligned} &\mathbf{P}(Tampering) \times \\ &\mathbf{P}(Fire) \times \\ &\mathbf{P}(Alarm \mid Tampering, Fire) \times \\ &\mathbf{P}(Smoke \mid Fire) \times \\ &\mathbf{P}(Leaving \mid Alarm) \times \\ &\mathbf{P}(Report \mid Leaving) \times \end{aligned}$$

The full joint probability table requires $2^6 - 1$ entries. Only 12 entries are needed for the conditional probabilities.

Fire Alarm Domain

Assume the following (conditional) probabilities (where we write $P(A \mid B)$ for $P(A = 1 \mid B = 1)$, $P(\neg A \mid B)$ for $P(A = 0 \mid B = 1)$ and so on):

- ▶ $P(\textit{Tampering}) = 0.02$
- ▶ $P(\textit{Fire}) = 0.01$
- ▶ $P(\textit{Smoke} \mid \textit{Fire}) = 0.9$
- ▶ $P(\textit{Smoke} \mid \neg \textit{Fire}) = 0.01$
- ▶ $P(\textit{Alarm} \mid \textit{Fire} \wedge \textit{Tampering}) = 0.5$
- ▶ $P(\textit{Alarm} \mid \textit{Fire} \wedge \neg \textit{Tampering}) = 0.99$
- ▶ $P(\textit{Alarm} \mid \neg \textit{Fire} \wedge \textit{Tampering}) = 0.85$
- ▶ $P(\textit{Alarm} \mid \neg \textit{Fire} \wedge \neg \textit{Tampering}) = 0.0001$
- ▶ $P(\textit{Leaving} \mid \textit{Alarm}) = 0.88$
- ▶ $P(\textit{Leaving} \mid \neg \textit{Alarm}) = 0.001$
- ▶ $P(\textit{Report} \mid \textit{Leaving}) = 0.75$
- ▶ $P(\textit{Report} \mid \neg \textit{Leaving}) = 0.01$

Querying

If Report is observed, then the probability of Fire and Tampering go up:

- ▶ $P(\text{Fire}) = 0.01$ and $P(\text{Fire} \mid \text{Report}) = 0.2305$
- ▶ $P(\text{Tampering}) = 0.02$ and $P(\text{Tampering} \mid \text{Report}) = 0.399$

If, in addition, Smoke is observed, then probability of Fire goes up further but Tampering goes down:

- ▶ $P(\text{Fire} \mid \text{Report} \wedge \text{Smoke}) = 0.964$
- ▶ $P(\text{Tampering} \mid \text{Report} \wedge \text{Smoke}) = 0.0284$

If, however, \neg Smoke is observed, then the probability of Fire goes down:

- ▶ $P(\text{Fire} \mid \text{Report} \wedge \neg \text{Smoke}) = 0.0294$
- ▶ $P(\text{Tampering} \mid \text{Report} \wedge \neg \text{Smoke}) = 0.501$

Summary Belief Networks

- ▶ Belief networks are a representation of **conditional independence** in probabilistic models.
- ▶ Querying can often be done using exact inference (with algorithmic tricks).
- ▶ Sometimes exact inference too hard. There are also approximate algorithms.
- ▶ Lots of research on **learning** belief networks from data. Either learning the conditional probabilities or even the structure of a belief network.

Expectation: motivation

- ▶ Consider a random variable F with values from the real numbers x_1, \dots, x_n .
- ▶ Assume $P(F = x_i) = 1/n$ for all x_i .
- ▶ Then the **expected value** $E[F]$ of F is the **average** over the possible values x_1, \dots, x_n of F :

$$\frac{x_1 + \dots + x_n}{n} = x_1 \frac{1}{n} + \dots + x_n \frac{1}{n} = \sum_{x_i} x_i P(F = x_i)$$

- ▶ If the probabilities $P(F = x_i)$ are not all equal, we take the probability-weighted average.

Expectation

Assume that x_1, \dots, x_n are the values a random variable F can take. Then the **expected value** of F is defined as follows:

$$E[F] = x_1P(F = x_1) + \dots + x_nP(F = x_n) = \sum_x xP(F = x)$$

In other words, $E[F]$ is the probability-weighted average of all possible values of F .

$E[F]$ is sometimes called the **expectation** of F or the **mean** of F .

Example

Suppose you roll a fair die.

To model this take a random variable F with values in

$$\{1, 2, 3, 4, 5, 6\}$$

and set $P(F = x) = 1/6$ for all $x \in \{1, 2, 3, 4, 5, 6\}$.

The expected value of the random variable F is

$$\begin{aligned} E[F] &= \sum_x xP(F = x) \\ &= 1P(F = 1) + 2P(F = 2) + 3P(F = 3) \\ &\quad + 4P(F = 4) + 5P(F = 5) + 6P(F = 6) \\ &= \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} \\ &= \frac{7}{2} \end{aligned}$$

Example

Suppose I pay, in p , the face value of a fair die if it comes up odd and earn the face value if it comes up even. What are my expectations?

- ▶ Consider the random variable F that takes values $\{-1, -3, -5, 2, 4, 6\}$ and $P(F = x) = 1/6$ for all $x \in \{-1, -3, -5, 2, 4, 6\}$.
- ▶ For example, $(F = -1)$ is the event that the face value is 1 and, thus, I pay $1p$.

The expected value of the random variable F is

$$\begin{aligned} E[F] &= \sum_x xP(F = x) \\ &= -P(F = -1) + 2P(F = 2) - 3P(F = -3) \\ &\quad + 4P(F = 4) - 5P(F = -5) + 6P(F = 6) \\ &= -\frac{1}{6} + \frac{2}{6} - \frac{3}{6} + \frac{4}{6} - \frac{5}{6} + \frac{6}{6} \\ &= \frac{1}{2} \end{aligned}$$

Another way to compute $E[F]$

Let F be a random variable from the probability space (S, P) .
Then

$$E[F] = \sum_{s \in S} F(s)P(s)$$

Proof. Assume F takes the values $\{x_1, \dots, x_n\}$. Then

$$\begin{aligned} \sum_{s \in S} F(s)P(s) &= \sum_{F(s)=x_1} F(s)P(s) + \cdots + \sum_{F(s)=x_n} F(s)P(s) \\ &= x_1 \sum_{F(s)=x_1} P(s) + \cdots + x_n \sum_{F(s)=x_n} P(s) \\ &= x_1 P(F = x_1) + \cdots + x_n P(F = x_n) \\ &= E[F] \end{aligned}$$

Linearity of Expectations

Let F and G be random variables from the same probability space (S, P) and let λ be a real number. Define new random variables $F + G$ and λF by setting:

$$(F + G)(s) = F(s) + G(s), \quad (\lambda F)(s) = \lambda F(s)$$

Then

$$E[F + G] = E[F] + E[G], \quad E[\lambda F] = \lambda E[F]$$

Proof. We prove the first claim using the reformulation of the expected value from the previous slide.

$$\begin{aligned} E[F + G] &= \sum_{s \in S} (F + G)(s) P(s) \\ &= \sum_{s \in S} (F(s) P(s) + G(s) P(s)) \\ &= \sum_{s \in S} F(s) P(s) + \sum_{s \in S} G(s) P(s) \\ &= E[F] + E[G] \end{aligned}$$

Example: Rolling two dice again

What is the expected value of the random variable F with $F(ab) = a + b$ from

$$S = \{1, 2, 3, 4, 5, 6\}^2$$

where $P(ab) = \frac{1}{36}$ for every $ab \in S$?

Note that $F(ab) = F_1(ab) + F_2(ab)$, where F_1 is the random variable from S with

$$F_1(ab) = a$$

and F_2 is the random variable from S with

$$F_2(ab) = b$$

We have

$$E(F) = E(F_1 + F_2) = E(F_1) + E(F_2) = \frac{7}{2} + \frac{7}{2} = 7$$

Expectation does not distribute over multiplication

If F and G are two random variables then, $E[F \times G]$ does not always equal $E[F] \times E[G]$.

Proof by counterexample. Consider the sample space (S, P) with

$$S = \{H, T\}, \quad P(H) = P(T) = \frac{1}{2}$$

Define random variables F_h and F_t by setting

$$F_h(H) = 1, \quad F_h(T) = 0, \quad F_t(H) = 0, \quad F_t(T) = 1$$

and define the random variable F by setting $F(s) = F_h(s) \times F_t(s)$.

Then $F(s) = F_h(s) \times F_t(s) = 0$ for all $s \in S$ and so $E[F] = 0$ but $E(F_h) \times E(F_t) = \frac{1}{4}$ since

$$E[F_h] = 1 \times \frac{1}{2} + 0 \times \frac{1}{2} = \frac{1}{2}, \quad E[F_t] = 0 \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{1}{2}$$

Expectation and Decision Making

Consider an agent who has to deliver a mail. There is a **short way** and a **long way** to deliver it. On the short way, it is more likely that the agent will have an **accident**. Suppose the agent can get **protectors** that will not change the probability of an accident but will make one less severe. The protectors are expensive, however. Going the long way reduces the probability of an accident, but takes much longer.

Problem: the agent has to decide whether to wear protectors and which way to go.

We model this situation using

- ▶ Two **decision variables**, *Which_way* and *Protector*. The agent gets to choose the value for each decision variable.
- ▶ A **random variable**, *Accident*, which represents whether an accident happens.
- ▶ A **utility function** that gives the utility of every possible outcome.

Conditional Probability Distribution

Recall that the probability of an accident only depends on whether the long or short way is chosen but not on whether the agent is wearing protectors. Thus, all we need is the conditional probability distribution

$$\mathbf{P}(\textit{Accident} \mid \textit{Which_way})$$

Assume this is given by

- ▶ $P(\textit{Accident} = 1 \mid \textit{Which_way} = \textit{short}) = 0.2$
- ▶ $P(\textit{Accident} = 0 \mid \textit{Which_way} = \textit{short}) = 0.8$
- ▶ $P(\textit{Accident} = 1 \mid \textit{Which_way} = \textit{long}) = 0.01$
- ▶ $P(\textit{Accident} = 0 \mid \textit{Which_way} = \textit{long}) = 0.99$

Utility

The **utility** of the outcome depends on whether a short or long way is chosen, whether the agent wears protectors, and whether the agent has an accident. Assume the utility values are as follows:

<i>Protector</i>	<i>Which_way</i>	<i>Accident</i>	<i>Utility</i>
<i>true</i>	<i>short</i>	<i>true</i>	35
<i>true</i>	<i>short</i>	<i>false</i>	95
<i>true</i>	<i>long</i>	<i>true</i>	30
<i>true</i>	<i>long</i>	<i>false</i>	75
<i>false</i>	<i>short</i>	<i>true</i>	3
<i>false</i>	<i>short</i>	<i>false</i>	100
<i>false</i>	<i>long</i>	<i>true</i>	0
<i>false</i>	<i>long</i>	<i>false</i>	80

What to do? Maximize Expected Utility!

Choose *Protector* and *Which_way* in such a way that the **expected utility** is maximal: choose values r, s such that

$$E[\text{Utility} \mid \text{Protector} = r, \text{Which_way} = s]$$

is maximal.

We compute $E[\text{Utility} \mid \text{Protector} = 1, \text{Which_way} = \text{short}]$ by taking the sum of

- ▶ the value **35** of the utility of $(\text{Protector} \wedge \text{short} \wedge \text{Accident})$ multiplied by $P(\text{Accident} = 1 \mid \text{Which_way} = \text{short})$;
- ▶ the value **95** of the utility of $(\text{Protector} \wedge \text{short} \wedge \neg \text{Accident})$ multiplied by $P(\text{Accident} = 0 \mid \text{Which_way} = \text{short})$.

We obtain:

$$\begin{aligned} 83 &= 35 \times P(\text{Accident} = 1 \mid \text{Which_way} = \text{short}) + \\ &\quad 95 \times P(\text{Accident} = 0 \mid \text{Which_way} = \text{short}) \end{aligned}$$

Maximize Expected Utility

$E[\text{Utility} \mid \text{Protector} = 1, \text{Which_way} = \text{long}]$ equals

$$\begin{aligned} 74.55 &= 30 \times P(\text{Accident} = 1 \mid \text{Which_way} = \text{long}) + \\ &\quad 75 \times P(\text{Accident} = 0 \mid \text{Which_way} = \text{long}) \end{aligned}$$

$E[\text{Utility} \mid \text{Protector} = 0, \text{Which_way} = \text{short}]$ equals

$$\begin{aligned} 80.6 &= 3 \times P(\text{Accident} = 1 \mid \text{Which_way} = \text{short}) + \\ &\quad 100 \times P(\text{Accident} = 0 \mid \text{Which_way} = \text{short}) \end{aligned}$$

$E[\text{Utility} \mid \text{Protector} = 0, \text{Which_way} = \text{long}]$ equals

$$\begin{aligned} 79.2 &= 0 \times P(\text{Accident} = 1 \mid \text{Which_way} = \text{long}) + \\ &\quad 80 \times P(\text{Accident} = 0 \mid \text{Which_way} = \text{long}) \end{aligned}$$

As $E[\text{Utility} \mid \text{Protector} = 1, \text{Which_way} = \text{short}]$ is maximal, a rational agent should wear protectors and take the short way.

Journey to Manchester Airport

Recall the following example from the introduction: When going to the airport by car, how early should I start? 45 minutes should be enough from Liverpool to Manchester Airport, but only under the assumption that there are no accidents, no lane closures, that my car does not break down, and so on. This uncertainty is hard to eliminate, but still an agent has to make a decision.

Within the framework of “maximizing expected utility” this problem can be modeled as follows.

Journey to Manchester Airport

Consider a random variable **Arrival** taking values:

- ▶ *miss_plane*,
- ▶ *wait_0_minutes*, *wait_5_minutes*, *wait_10_minutes*, and so on

Consider the decision variable **Start** taking values:

- ▶ *45minutes_early*, *50minutes_early*, and so on.

Assume from experience/data we have a probability distribution

$$\mathbf{P}[\textit{Arrival} \mid \textit{Start}]$$

Finally, consider **utility values** associated with the values of the random variables **Arrival**: *miss_plane* has very low utility, *wait_0_minutes* high utility, and so on.

Then the agent should choose the value r for **Start** such that the expected utility

$$E[\textit{Utility} \mid \textit{Start} = r]$$

is maximal.