5. Reasoning Under Uncertainty

Contents

Handling Uncertain Knowledge

- learned knowledge representation using first-order logic and propositional logic with certainty
- sure about the predicates.
- With this knowledge representation, we might write A→B, which means if A is true then B is true,
- but consider a situation where we are not sure about whether A is true or not then we cannot express this statement, this situation is called uncertainty.
- Hence need uncertain reasoning or probabilistic reasoning.

Causes of uncertainty

1. Information occurred from unreliable sources.
2. Experimental Errors
3. Equipment fault
4. Temperature variation
5. Climate change.

Probabilistic Reasoning

- way of knowledge representation where we apply the concept of probability to indicate the uncertainty in knowledge
- In probabilistic reasoning, we combine probability theory with logic to handle the uncertainty
- E.g:
  - "It will rain today,"
  - "behavior of someone for some situations,"
  - "A match between two teams or two players."

These are probable sentences for which we can assume that it will happen but not sure about it, so here we use probabilistic reasoning.

Need of Probabilistic Reasoning in AI

- When there are unpredictable outcomes.
- When an unknown error occurs during an experiment.

In probabilistic reasoning, there are two ways to solve problems with uncertain knowledge:

- **Bayes' rule**
- **Bayesian Statistics**

Probability

- Probability can be defined as a chance that an uncertain event will occur.
- It is the numerical measure of the likelihood that an event will occur.
- The value of probability always remains between 0 and 1 that represent ideal uncertainties.
    a. $0 \leq P(A) \leq 1$, where P(A) is the probability of an event A.
    b. $P(A) = 0$, indicates total uncertainty in an event A.
    c. $P(A) = 1$, indicates total certainty in an event A.
- the probability of an uncertain event by using the below formula:

$$\textbf{Probability of occurrence} = \frac{\text{Number of desired outcomes}}{\text{Total number of outcomes}}$$

Probability

- P(¬A) = probability of a not happening event.
- P(¬A) + P(A) = 1.

Event: Each possible outcome of a variable is called an event.

Sample space: The collection of all possible events is called sample space.

Random variables: Random variables are used to represent the events and objects in the real world.

Prior probability: The prior probability of an event is probability computed before observing new information.

Posterior Probability: The probability that is calculated after all evidence or information has taken into account. It is a combination of prior probability and new information.

Conditional Probability

- Conditional probability is a probability of occurring an event when another event has already happened.
- Let's suppose, we want to calculate the event A when event B has already occurred, "the probability of A under the conditions of B", it can be written as:

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}$$

- **Where P($A \wedge B$)= Joint probability of a and B**
- **P(B)= Marginal probability of B.**

Conditional Probability

If the probability of A is given and we need to find the probability of B, then it will be given as:

$$P(B|A) = \frac{P(A \wedge B)}{P(A)}$$

Example

In a class, there are 70% of the students who like English and 40% of the students who likes English and mathematics, and then what is the percent of students those who like English also like mathematics?

Solution:

Let, A is an event that a student likes Mathematics

B is an event that a student likes English.

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} = \frac{0.4}{0.7} = 57\%$$

Hence, 57% are the students who like English also like Mathematics.

Example

In a survey among a group of students, 70% play football, 60% play basketball, and 40% play both sports. If a student is chosen at random and it is known that the student plays basketball, what is the probability that the student also plays football?

Solution

Let's assume there are 100 students in the survey.

Number of students who play football = n(A) = 70

Number of students who play basketball = n(B) = 60

Number of students who play both sports = n(A ∩ B) = 40

*To find the probability that a student plays football given that they play basketball, we use the **conditional probability formula**:*

P(A|B) = n(A ∩ B) / n(B)

Substituting the values, we get:

P(A|B) = 40 / 60 = 2/3

Therefore, the probability that a randomly chosen student who plays basketball also plays football is 2/3.

Bayes' Theorem

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

P(A|B) is known as **posterior**, which we need to calculate, and it will be read as Probability of hypothesis A when we have occurred an evidence B.

P(B|A) is called the likelihood, in which we consider that hypothesis is true, then we calculate the probability of evidence.

P(A) is called the **prior probability**, probability of hypothesis before considering the evidence

P(B) is called **marginal probability**, pure probability of an evidence.

Example 1

Question: what is the probability that a patient has diseases meningitis with a stiff neck?

Given Data:

A doctor is aware that disease meningitis causes a patient to have a stiff neck, and it occurs 80% of the time. He is also aware of some more facts, which are given as follows:

- The Known probability that a patient has meningitis disease is 1/30,000.
- The Known probability that a patient has a stiff neck is 2%.

Solution 1

Let a be the proposition that patient has stiff neck and b be the proposition that patient has meningitis. , so we can calculate the following as:

P(a|b) = 0.8

P(b) = 1/30000

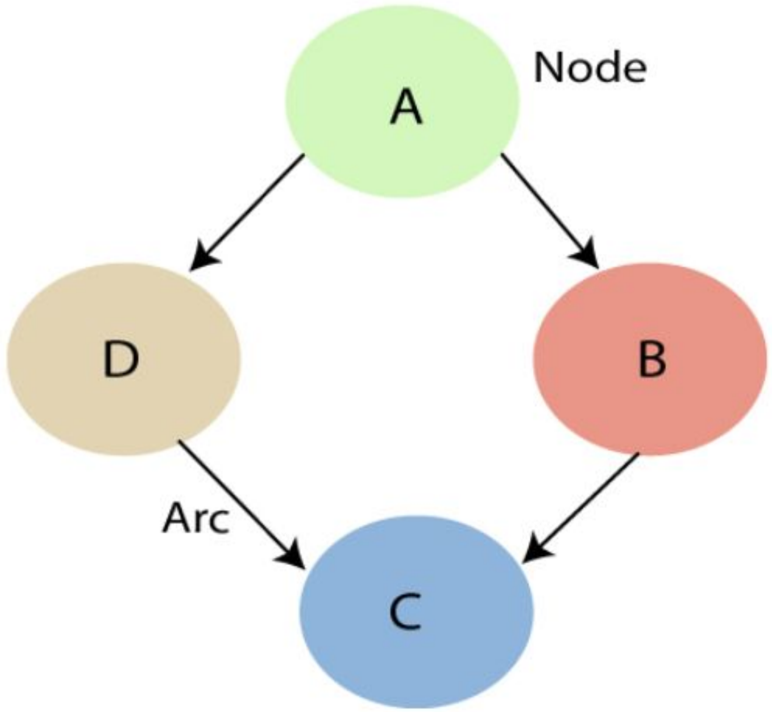
P(a)= .02

$$\mathbf{P(b\,|\,a)} = \frac{P(a|b)P(b)}{P(a)} = \frac{0.8*(\frac{1}{30000})}{0.02} = 0.001333333.$$

Bayesian Belief Network

- "A Bayesian network is a probabilistic graphical model which represents a set of variables and their conditional dependencies using a directed acyclic graph."
- It is also called a **Bayes network, belief network, decision network**, or **Bayesian model**.
- it consists of two parts:
  - **Directed Acyclic Graph**
  - **Table of conditional probabilities.**
- **A Bayesian network graph is made up of nodes and Arcs (directed links), where:**
  - Each **node** corresponds to the random variables
  - a variable can be **continuous** or **discrete**.
  - **Arc or directed arrows** represent conditional probabilities between random variables. These directed links or arrows connect the pair of nodes in the graph.
  - These links represent that one node directly influence the other node
  - if there is no directed link that means that nodes are independent with each other

Bayesian Belief Network



- A, B, C, and D are random variables represented by the nodes of the network graph.
- If we are considering node B, which is connected with node A by a directed arrow, then node A is called the parent of Node B.
- Node C is independent of node A.

Joint Probability Distribution

- Bayesian network is based on Joint probability distribution and conditional probability.
- If we have variables x1, x2, x3,....., xn, then the probabilities of a different combination of x1, x2, x3.. xn, are known as Joint probability distribution.
- $\mathbf{P[x_1, x_2, x_3, \ldots, x_n]}$, it can be written as the following way in terms of the joint probability distribution.

$$= \mathbf{P[x_1 | x_2, x_3, \ldots, x_n] P[x_2, x_3, \ldots, x_n]}$$

$$P(X_i | X_{i-1}, \ldots, X_1) = P(X_i \mid Parents(X_i))$$

Example: 1(Inference using full joint distribution)

(b) Find the probabilistic inference by enumeration of entries in a full joint distribution table shown in figure 1. **10**

(i) No cavity when toothache is there

(ii) p (Cavity! toothache or catch)

	toothache		¬toothache	
	catch	¬catch	catch	¬catch
cavity	**.108**	**.012**	**.072**	**.008**
¬cavity	**.016**	**.064**	**.144**	**.576**

Example: 1: Solution

1. **No cavity when toothache is there**

$$P(\neg cavity|toothache) = P(\neg cavity^t oothache)/P(toothache)$$
$$= (0.016 + 0.064)/(0.108 + 0.012 + 0.016 + 0.064)$$
$$= 0.4$$

2. **p(Cavity|toothache or catch)**

$$= P(Cavity|toothache) + p(catch)$$
$$= [P(cavity\hat{t}oothache)/P(toothache)] + P(catch)$$
$$= [(0.108 + 0.012)]$$
$$/[(0.108 + 0.012 + 0.016 + 0.064 + (0.108 + 0.016 + 0.072 + 0.144)$$
$$= 0.96$$

Example: 2(Inference using full joint distribution)

A men's soccer team plays soccer zero, one, or two days a week. The probability that they play zero days is .2, the probability that they play one day is .5, and the probability that they play two days is .3. Find the long-term average or expected value, μ, of the number of days per week the men's soccer team plays soccer.

To find the expected value, *E(X)*, or mean μ of a discrete random variable *X*, simply multiply each value of the random variable by its probability and add the products. The formula is given as $E(X) = \mu = \sum xP(x)$.

Example: 2(Inference using full joint distribution): Solution

Expected Value Table

x	$P(x)$	$x^* P(x)$
0	.2	(0)(.2) = 0
1	.5	(1)(.5) = .5
2	.3	(2)(.3) = .6

Add the last column $x * P(x)$ to get the expected value/mean of the random variable *X*.

$$E(X) = \mu = \sum xP(x) = 0 + .5 + .6 = 1.1$$

The expected value/mean is 1.1. The men's soccer team would, on the average, expect to play soccer 1.1 days per week.

Example: 3(Inference using full joint distribution)

The joint probability function of two discrete r.v's X and Y is given by $f(x, y) = c(2x+y)$, where x and y can assume all integers such that $0 \leq x \leq 2, 0 \leq y \leq 3$ and $f(x,y) = 0$ otherwise. Find $E(X)$, $E(Y)$

Example: 3(Inference using full joint distribution): Solution

To find c: We can tabulate the probabilities as follows:

$f(x,y) = c(2x+y) \quad 0 \le x \le 2, 0 \le y \le 3$

=0

$X \diagdown Y$	**0**	**1**	**2**	**3**	**Total**
0	0	c	2c	3c	6c
1	2c	3c	4c	5c	14c
2	4c	5c	6c	7c	22c
Total	6c	9c	12c	15c	42c

Since $\sum p_i = 1$

$\therefore$ 42c=1

$\therefore c = \frac{1}{42}$

Example: 3(Inference using full joint distribution): Solution

With this value the probability distribution is

$X \diagdown Y$	0	1	2	3	Total
0	0	$\frac{1}{42}$	$\frac{2}{42}$	$\frac{3}{42}$	$\frac{6}{42}$
1	$\frac{2}{42}$	$\frac{3}{42}$	$\frac{4}{42}$	$\frac{5}{42}$	$\frac{14}{42}$
2	$\frac{4}{42}$	$\frac{5}{42}$	$\frac{6}{42}$	$\frac{7}{42}$	$\frac{22}{42}$
Total	$\frac{6}{42}$	$\frac{9}{42}$	$\frac{12}{42}$	$\frac{15}{42}$	1

Example: 3(Inference using full joint distribution): Solution

∴ The marginal probability distributions of X & Y are:

X	P(X)
0	$\frac{6}{42}$
1	$\frac{14}{42}$
2	$\frac{22}{42}$

X	P(Y)
0	$\frac{6}{42}$
1	$\frac{9}{42}$
2	$\frac{12}{42}$
3	$\frac{15}{42}$

Example: 3(Inference using full joint distribution): Solution

$$E(X) = \sum p_i x_i$$

$$=0+1\times\frac{14}{42} + 2 \times \frac{22}{42}$$

$$E(X) = \frac{58}{42} = 1.381$$

$$E(Y) = \sum p_i y_i$$

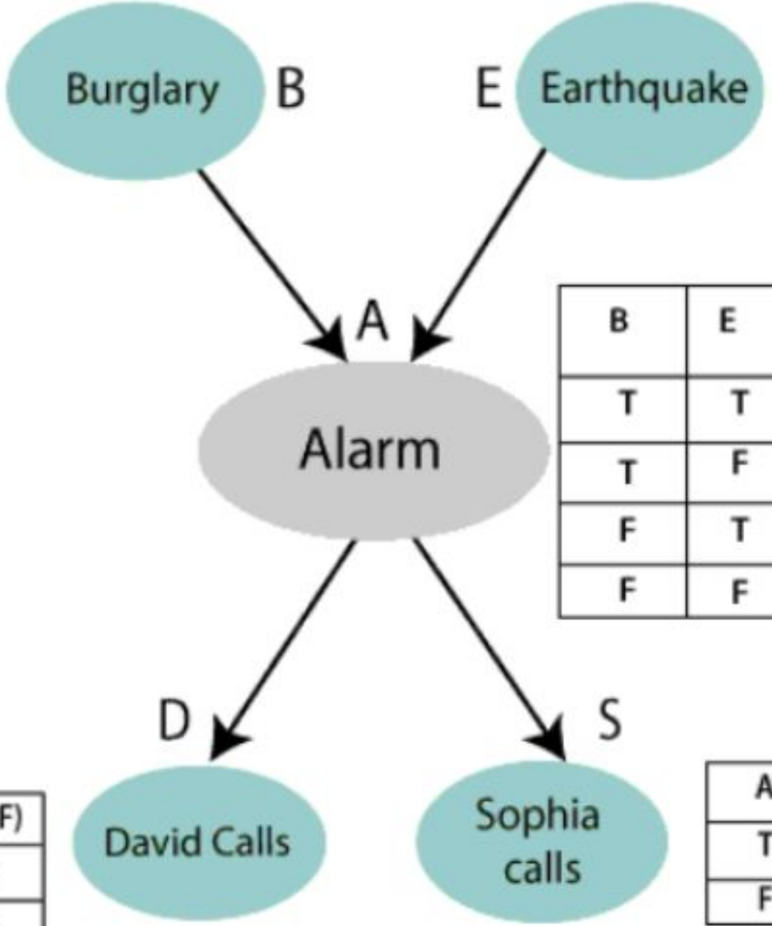
$$= 0 + 1 \times \frac{9}{42} + 2 \times \frac{12}{42} + 3 \times \frac{15}{42}$$

$$E(Y) = \frac{78}{42} = 1.857$$

Example1 Reasoning in Belief Network

Harry installed a new burglar alarm at his home to detect burglary. The alarm reliably responds at detecting a burglary but also responds for minor earthquakes. Harry has two neighbors David and Sophia, who have taken a responsibility to inform Harry at work when they hear the alarm. David always calls Harry when he hears the alarm, but sometimes he got confused with the phone ringing and calls at that time too. On the other hand, Sophia likes to listen to high music, so sometimes she misses to hear the alarm. Here we would like to compute the probability of Burglary Alarm.

Calculate the probability that alarm has sounded, but there is neither a burglary, nor an earthquake occurred, and David and Sophia both called the Harry.



T	0.002
F	0.998

T	0.001
F	0.999

B	E	P(A=T)	P(A=F)
T	T	0.94	0.06
T	F	0.95	0.04
F	T	0.31	0.69
F	F	0.0001	0.999

A	P (D=T)	P (D=F)
T	0.91	0.09
F	0.05	0.95

A	P (S=T)	P (S=F)
T	0.75	0.25
F	0.02	0.98

Example:1: Solution

List of all events occurring in this network:

- **Burglary (B)**
- **Earthquake(E)**
- **Alarm(A)**
- **David Calls(D)**
- **Sophia calls(S)**

We can write the events in the form of probability: **P[D, S, A, B, E]**

Example:1: Solution

The above probability statement using joint probability distribution:

P[D, S, A, B, E]= P[D | S, A, B, E]. P[S, A, B, E]

=P[D | S, A, B, E]. P[S | A, B, E]. P[A, B, E]

= P [D| A]. P [S| A, B, E]. P[A, B, E]

= P[D | A]. P[S | A]. P[A| B, E]. P[B, E]

= P[D | A]. P[S | A]. P[A| B, E]. P[B]. P[E]

Example:1: Solution

Let's take the observed probability for the Burglary and earthquake component:

P(B= True) = 0.002, which is the probability of burglary.

P(B= False)= 0.998, which is the probability of no burglary.

P(E= True)= 0.001, which is the probability of a minor earthquake

P(E= False)= 0.999, Which is the probability that an earthquake not occurred.

Example:1: Solution

conditional probabilities as per the below tables:

B	E	P(A= True)	P(A= False)
True	True	0.94	0.06
True	False	0.95	0.04
False	True	0.31	0.69
False	False	0.001	0.999

Example:1: Solution

The Conditional probability of David that he will call depends on the probability of Alarm.

A	P(D= True)	P(D= False)
True	0.91	0.09
False	0.05	0.95

Example:1: Solution

The Conditional probability of Sophia that she calls is depending on its Parent Node "Alarm."

A	P(S= True)	P(S= False)
True	0.75	0.25
False	0.02	0.9

Example:1: Solution

From the formula of joint distribution, we can write the problem statement in the form of probability distribution:

P(S, D, A, ¬B, ¬E) = P (S|A) *P (D|A)*P (A|¬B ^ ¬E) *P (¬B) *P (¬E).

= 0.75* 0.91* 0.001* 0.998*0.999

= 0.00068045.

Example 2: Reasoning in Belief Network

It is known that whether or not a person has cancer is directly influenced by whether she is exposed to second-hand smoke and whether she smokes. Both of these things are affected by whether her parents smoke. Cancer reduces a person's life expectancy.

(i)Draw the Bayesian Belief Network for the above situation.

(ii)Associate a conditional probability table for each node.

Example 2: Solution

(i)Draw the Bayesian Belief Network for the above situation.

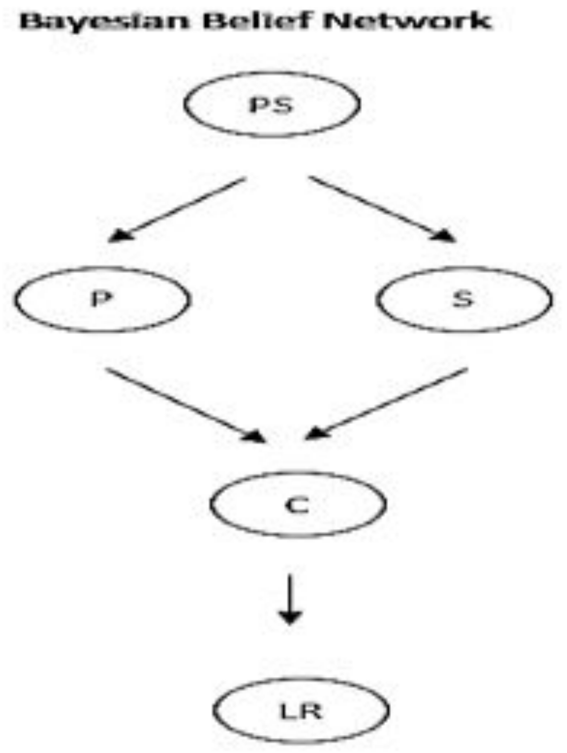
PS – Person’s Parents Smoke

P – Person Smokes

S – Person exposed to Second-hand Smoke

C – Person has Cancer

LR – Person’s Life Expectancy is reduced



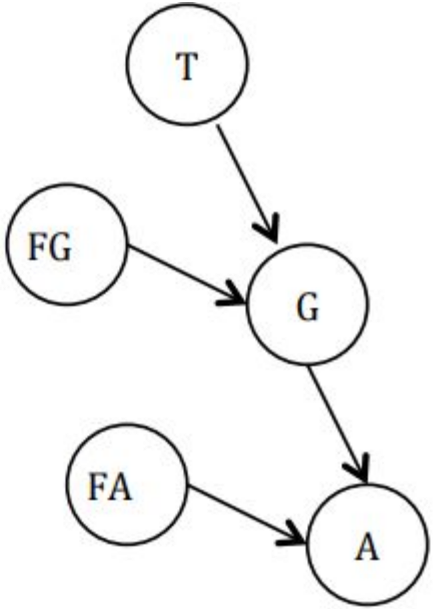
Example 3

The gauge reading at a nuclear power station shows high values if the temperature of the core goes very high. The gauge also shows high value if the gauge is faulty. A high reading in the gauge sets an alarm off. The alarm can also go off if it is faulty. The probability of faulty instruments is low in a nuclear power plant.

i) Draw the Bayesian Belief Network for the above situation

ii) Associate a conditional probability table for each node.

Example 3: Solution

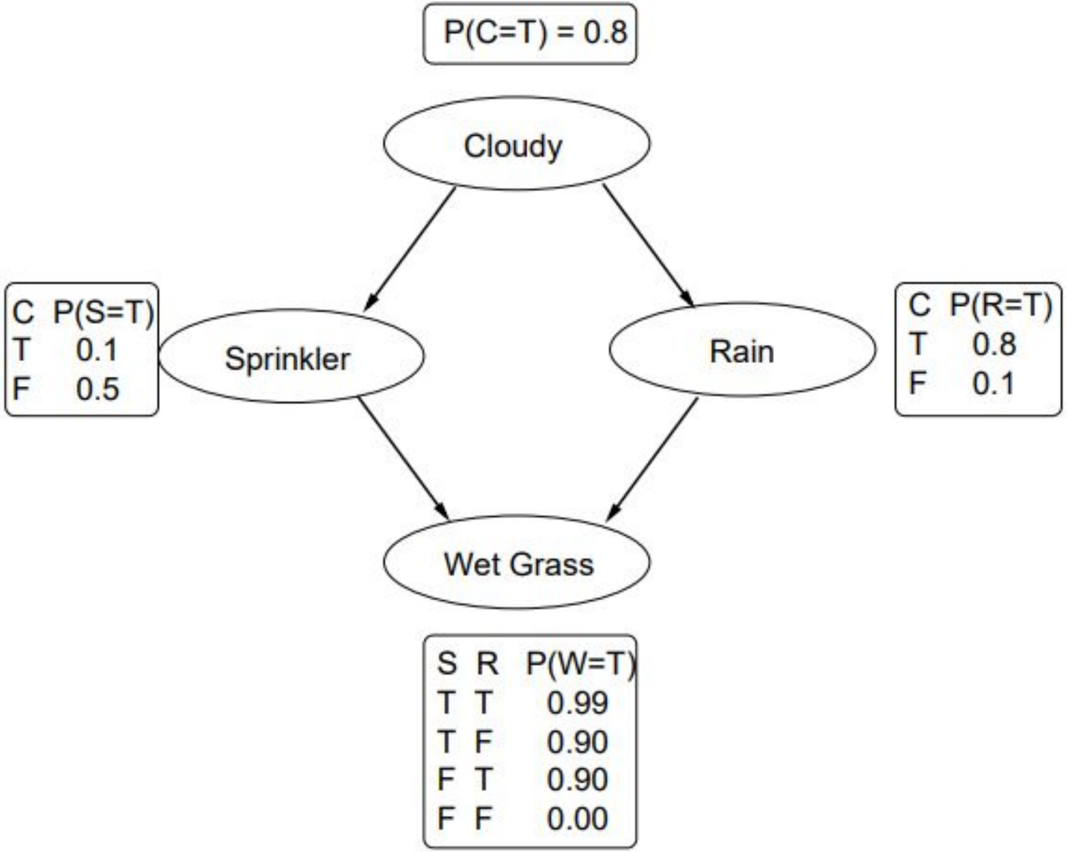


Example 4

Whether the grass is wet, W, depends on whether the sprinkler has been used, S, or whether it has rained, R. Whether the sprinkler is used depends on whether it is cloudy, similarly for whether it has rained. The probability of the grass being wet is conditionally independent of it being cloudy, given information about the sprinklers and whether it has rained.

“It is cloudy, what’s the probability that the grass is wet?”

Example 4: Solution



Example 4: Solution

$$P(W_{\mathrm{T}}|C_{\mathrm{T}}) = \frac{P(W_{\mathrm{T}}, C_{\mathrm{T}})}{P(C_{\mathrm{T}})}$$

The denominator is known (0.8). The numerator may expressed as a marginal distribution

$$\begin{aligned} P(W_{\mathrm{T}}, C_{\mathrm{T}}) &= \sum_S \sum_R P(W_{\mathrm{T}}, S, R, C_{\mathrm{T}}) \\ &= \sum_S \sum_R P(W_{\mathrm{T}}|S, R) P(S|C_{\mathrm{T}}) P(R|C_{\mathrm{T}}) P(C_{\mathrm{T}}) \end{aligned}$$

where the summation are over the variable being `T`, or `F`. From the simple example this is (note $P(C_{\mathrm{T}})$ has simply been cancelled from the numerator and denominator)

$$\begin{aligned} P(W_{\mathrm{T}}|C_{\mathrm{T}}) &= 0.99 \times 0.1 \times 0.8 + 0.90 \times 0.1 \times 0.2 \\ &\quad +0.90 \times 0.9 \times 0.8 + 0.00 \times 0.9 \times 0.2 \\ &= 0.7452 \end{aligned}$$