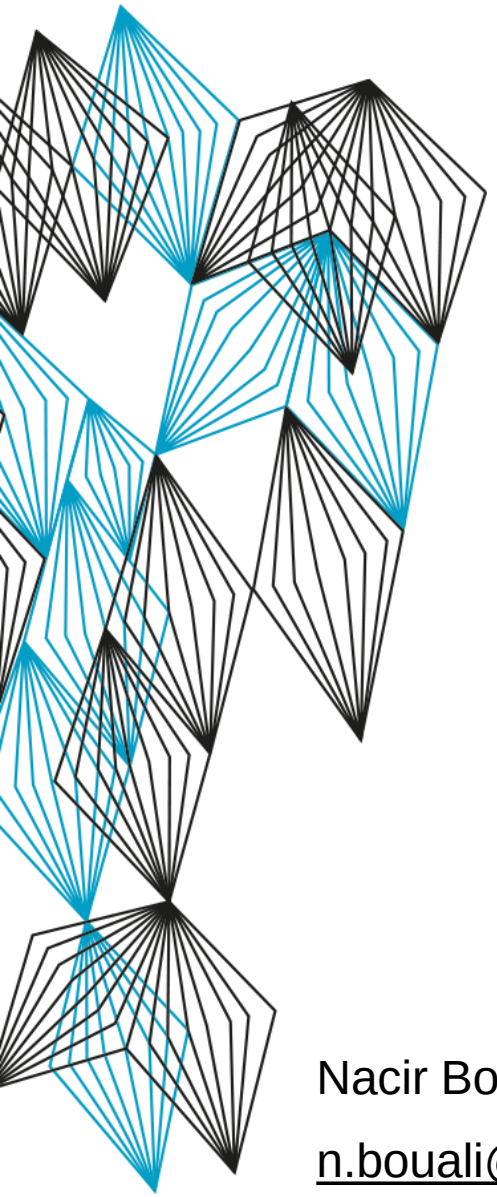


ARTIFICIAL INTELLIGENCE & CYBER SECURITY

EXAMPLE: BACK TO ROBOCOP

Nacir Bouali

n.bouali@utwente.nl

UNIVERSITY OF TWENTE.

ROBOCOP EXAMPLE

- Robocop is walking around in the city at night and there is car accident. A witness of the nighttime accident involving the car declares that the car was green.
- Do we have:
  - wb ⇒ b?
  - wg ⇒ g?
  - b ⇒ wb?
  - g ⇒ wg?
- Question: Is the car green or not?
  - Cannot be answered in a true/false way.
- But maybe we can compute the probability that the car is indeed blue

wb means "Witness declares car is blue".
b means "Car is actually blue".

ROBOCOP EXAMPLE

- In real life, the situation is different.
  - Robocup knows:
    - That cars in town are either blue or green.
    - It is known that under dim lighting conditions discrimination by humans between blue and green is 75% reliable;
      which means that $P(W = wb \mid C = b) = 0.75$ $P(W = wg \mid C = g) = 0.75$
    - That 8 out of 10 cars are actually blue. So without any witness a blue car is more likely.

ROBOCOP EXAMPLE

- Given the declaration of our witness should Robocop look for a green or blue car? In other words: what is the most likely color of the car?
- What is the optimal/rational decision under uncertainty?
- How do we model this, and can we reason about this?

ROBOCOP EXAMPLE

- What does Robocop want to know:
  - P(C = g | W = wg), or P(g | wg) for short, the probability that the car is green given that the witness stated that the car was green.
  - P(b | wg) the probability that the car is blue given that the witness stated that the car was green.
- How to derive/compute these probabilities?

ROBOCOP EXAMPLE

- Robocop can derive the following joint probabilities from the information it has.
- $P(b) = 0.8, P(wb|b) = 0.75, P(wg|g) = 0.75$

Joint Probability

	wg	wb	
b	0.2	0.6	0.8
g	0.15	0.05	0.2
	0.35	0.65	1

ROBOCOP EXAMPLE

- Robocop can derive the following joint probabilities from the information it has.
- $P(b) = 0.8, P(wb|b) = 0.75, P(wg|g) = 0.75$
- Using the Bayes rule, we can compute the probability the the car is green given that the witness said the car is green.
- $P(g|wg) = P(wg|g) * \frac{P(g)}{P(wg)}$

- $P(g|wg)$ =0.43 meaning that $P(b|wg) = 0.57$
- The rational decision is to look for a blue car.

Joint Probability

	wg	wb	
b	0.2	0.6	0.8
g	0.15	0.05	0.2
	0.35	0.65	1

URNS EXAMPLE

- Given an urn H with 5 white balls and 3 blue balls, let Xw (Xb) be the number of white (blue) balls drawn. One draws the balls with replacement.
- What is $P(X_w = 3, X_b = 2)$?
- Some considerations:
  - Since we replace the balls, the past does not affect future draws:
    - in all cases P(w) = 5/8 and P(b) = 3/8
  - We can draw 3 white and 2 blue balls in a number of ways:

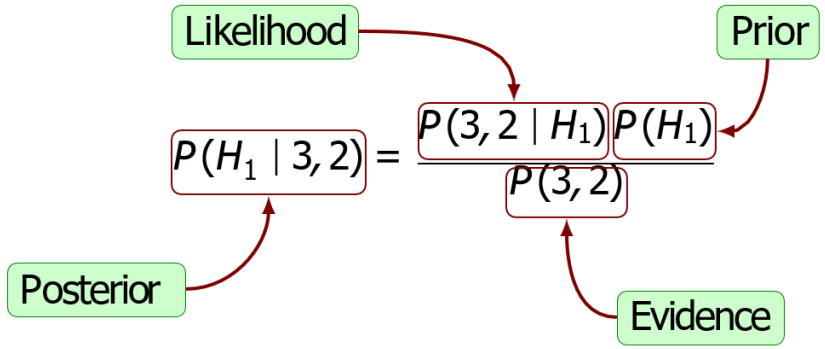
{w,w,w, b, b}; {w,w, b,w, b}; {w,w, b, b,w}; {w, b,w,w, b}; {w, b,w, b,w};…, in total $\binom{5}{3}$ 5 choose 3 = 10 ways

$$P(X_w = 3, X_b = 2) = \binom{5}{3}\left[\frac{5}{8}\right]^3\left[\frac{3}{8}\right]^2 \approx 0.34$$

URNS EXAMPLE

Given two urns H_1 and H_2, H_1 contains 5 white balls and 3 blue balls, urn H_2 contains 3 white balls and 5 blue balls. Assume that one randomly selects an urn and then draws 5 balls with replacement. Let the outcome of this chance experiment be $(X_w = 3, X_b = 2)$.

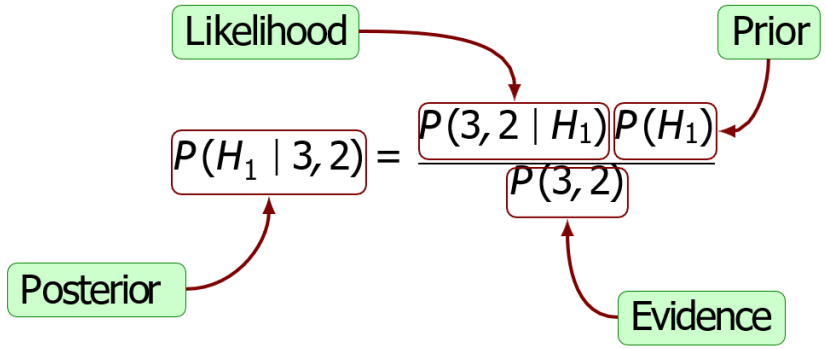
- Which urn was most likely selected, H_1 or H_2?
- One has to compute $P(H_1 \mid 3,2)$ and $P(H_2 \mid 3,2)$
- If $P(H_1 \mid 3,2) > P(H_2 \mid 3,2)$ then most likely H_1 was selected; otherwise H_2
- How do we compute $P(H_1 \mid 3,2)$?



UNIVERSITY OF TWENTE.

URNS EXAMPLE

Given two urns H_1 and H_2, H_1 contains 5 white balls and 3 blue balls, urn H_2 contains 3 white balls and 5 blue balls. Assume that one randomly selects an urn and then draws 5 balls with replacement. Let the outcome of this chance experiment be $(X_w = 3, X_b = 2)$.

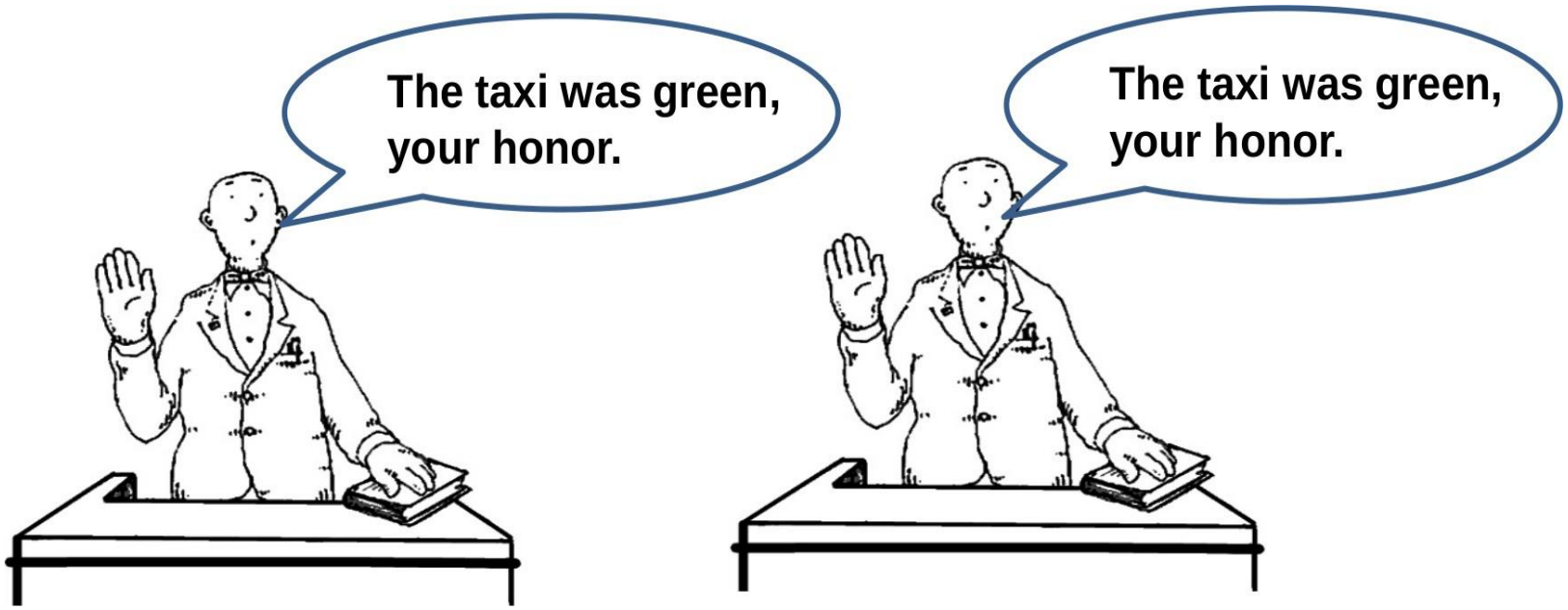


- $P(3,2 \mid H_1) = 0.3433$ and $P(H_1) = 0.5$
  , so $P(3,2 \mid H_1)\,P(H_1) = 0.1717$
- $P(3,2 \mid H_2) = \binom{5}{3}\left(\frac{3}{8}\right)^3\left(\frac{5}{8}\right)^2 = 0.20$
  and $P(H_2) = 0.5$, so $P(3,2 \mid H_2)\,P(H_2) = 0.1030$

Calculation P(3,2) doesn't change the result, but allows both values to sum to 1.
In our case, $\alpha = 0.1717 + 0.1030 = 0.2747$
P(H1|3,2)=0.1717/0.2747=0.625 , P(H2|3,2)= 0.375

BACK TO ROBOCOP

- What if another witness enters the scene and declares, independently from the first witness, that the car is green?
- Calculate $P(g \mid wg_1, wg_2)$



ROBOCOP EXAMPLE

- Independent observations are not directly affected by each other, but only by the state of the world (in this case, the actual colour of the car)

$$P(wg_1|g, wg_2) = P(wg_1|g) \quad \text{and} \quad P(wg_2|g, wg_1) = P(wg_2|g)$$

- It does not mean that knowledge of the statement of witness 1 would not give us any information about what witness 2 will declare (and vice versa).

On the contrary:

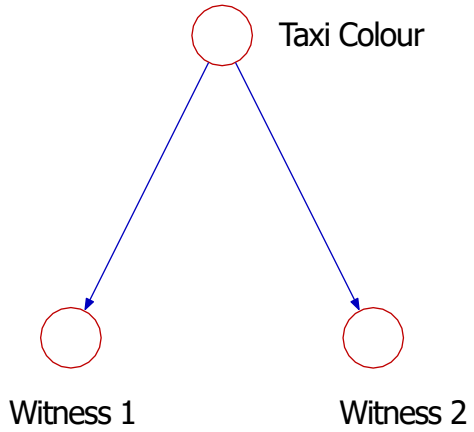
- Knowledge of the statement of witness 1 does give us information about what witness 2 will declare (and vice versa), if we do not know the color of the taxi.

MODELLING INDEPENDENCE

- Independent observations are not directly affected by each other, but only by the state of the world (in this case, the actual colour of the car)

$$P(wg_1|g, wg_2) = P(wg_1|g) \quad \text{and} \quad P(wg_2|g, wg_1) = P(wg_2|g)$$

$$P(A_1, \ldots, A_n) = P(A_1)P(A_2 \mid A_1)P(A_3 \mid A_1, A_2)\ldots P(A_n \mid A_1, \ldots, A_{n-1})$$
$$= \prod_{i=1}^{n} P(A_i \mid A_1, \ldots, A_{i-1})$$



$$P(W1, W2, W3, C) = P(W1|W2, W3, C)P(W2, W3, C) =$$
$$P(W1|W2, W3, C)P(W2|W3, C)P(W3, C) =$$
$$P(W1|W2, W3, C)P(W2|W3, C)\mathrm{P(W3|C)P(C)}$$

UNIVERSITY OF TWENTE.

PROSECUTOR'S FALLACY

"The chances of finding this evidence in an innocent man are so small that you can safely disregard the possibility that this man is innocent"

$P(E|\text{Innocent})$ is small

$\not\Rightarrow P(\text{Innocent}|E)$ is small

$\Rightarrow$ GUILTY!

CONCLUSION

- We have reviewed some basic concepts in probability theory.
- We have moved from the used of joint probabilities to working with conditional probability tables.
- We have also seen that Bayesian networks can make knowledge representation and inference more efficient.