Fundamentals of AI and KR - Module 3

6. Exercises & wrap-up

Paolo Torroni

Fall 2024

Notice

Downloading and ~~sharing~~

A copy of these slides can be downloaded from virtuale and stored for personal use only. Please do not redistribute.

Simple case studies

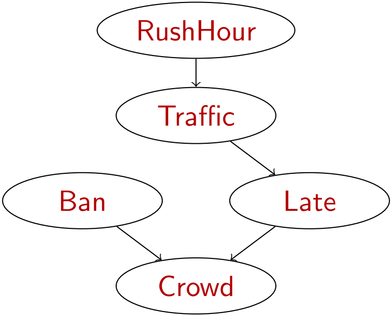
Bayes stop

You've been waiting at the bus stop for a while. While you're there, you realize that there's a small crowd of people waiting. Why is that so? You start thinking...

- a crowd at the bus stop is usually formed when the bus is delayed, or when more people than usual take the bus;
- heavy traffic often causes delays;
- there's almost always heavy traffic on rush hours;
- more people tend to use the bus when there's a car ban against pollution, which happens twice a week.

You open your notepad and sketch the following Bayesian network.

Bayes stop



All variables are Boolean:

- RushHour: it's rush hour (possible values: RushHour=r, RushHour=¬r);
- Traffic: there's heavy traffic (t/¬t);
- Late: the bus is delayed (l/¬l);
- Ban: a car ban is imposed (b/¬b);
- Crowd: there's a crowd at the bus stop (c/¬c).

While you're totally absorbed in drawing your sketch, you loose your sense of time, and you can no longer tell the day of the week, or the time of day.

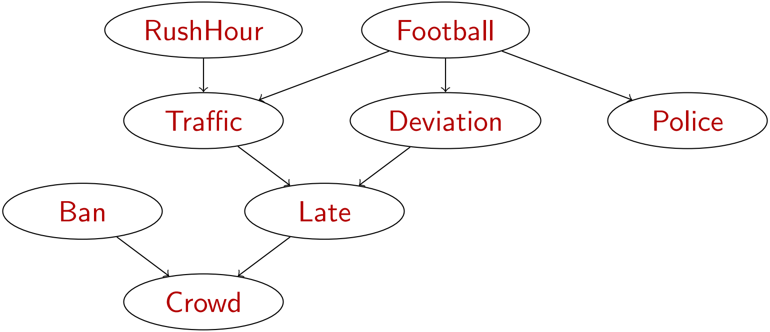
Bayes stop

1. Define reasonable CPTs for your Bayesian network. How many parameters are independent?
2. Can you use the network to explain the unusual crowd? Write down a probability query to show what kind of reasoning you can perform.
3. You ask around. Nobody knows if a car ban has been imposed. However, you learn that it's past the rush hour. Does that affect your belief about a possible car ban? How?

Bayes stop

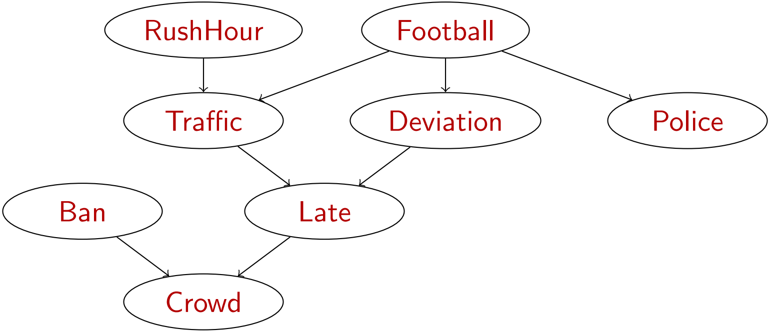
4. Somebody mentions a possible football match scheduled for tonight – which gets you thinking: before a match, traffic usually gets congested, and busses are deviated onto longer routes, which also contributes to delays. When there is a football match usually there are also some police patrols parked nearby.
   - Extend the Bayesian network to take these new aspects into account (Football, Deviation, Police).
   - Is the resulting network a polytree?

Bayes stop



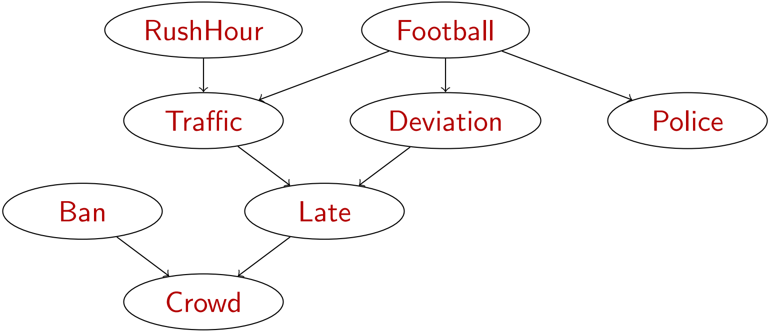
5. Is it true that P $\models$ (Police $\perp$ Ban $|$ Crowd, Deviation, RushHour)?
6. What is Deviation's Markov blanket?

Bayes stop



7. Consider the query: P(Traffic|RushHour=False,Deviation=True). How can you evaluate it using variable elimination? Show only the first step: $P(T|\neg r, d) = \alpha \ldots$

Bayes stop



8. Consider the query: P(Traffic|RushHour=False,Deviation=True). Generate one sample using a sampling method of your choice, and the following sequence of random numbers: 0.32, 0.92, 0.02, 0.05, 0.83, 0.59, 0.77, . . .

Bookworms

Tinker, tailor, soldier, spy! Your book collection contains thousands of volumes, mostly spy fiction novels, but also a few cookbooks, and some artificial intelligence textbooks. You wish to organize them by genre. However, you don't have the time to read them all. And maybe, you don't need to, either ... of course, you can use a Bayesian network!

The idea is simple: the **words** written in a **book** can tell you a lot about the book's **genre**. So, if you open a book at a random page and you see words like *enemy*, *agent*, or *Langley*, that could indicate that you're more likely to be holding a spy novel rather than a recipe book or an AI textbook. Excited by this idea, you decide to build a small-scale prototype, with a handful of selected words, and see where this gets you.

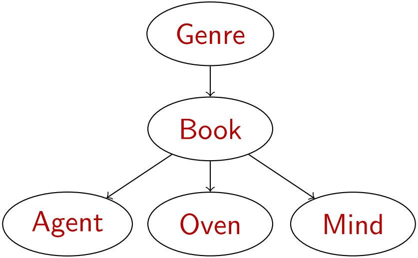
Bookworms

You begin to draw the following nodes:

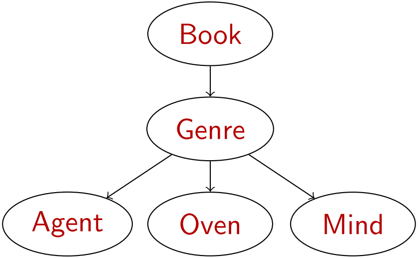
- Book: the particular book you pick from the shelf (it could be a unique integer number representing the book)
- Genre: the book's genre (spy, cooking, or ai);
- Agent: the word *agent* is present in a random page of the book.
- Oven: the word *oven* is present in a random page of the book.
- Mind: the word *mind* is present in a random page of the book.

However, you are not sure how to connect the nodes. Here are two options:

Bookworms



(a)



(b)

Bookworms

1. Which option can more directly express the probabilities of words in genres and of genres in books?
2. Consider both options (a) and (b): are Agent and Mind independent from one another? How about if you know the Genre?
3. Consider option (a), and assume that you know which particular book you picked – for example, book number 482. Can the probabilities of words tell you anything about the book's genre?

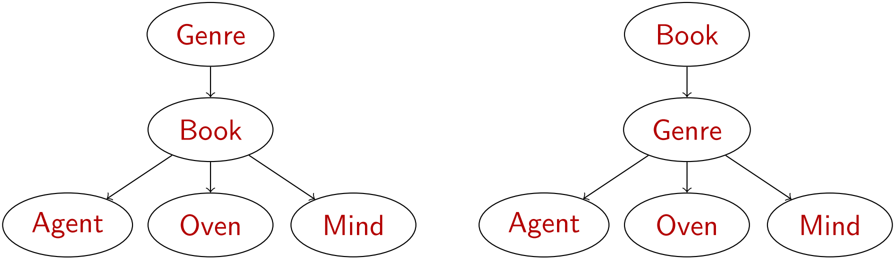
Bookworms

4. Based on the above considerations, which option better serves the purpose of estimating a book's genre by looking at random words, (a) or (b)?
5. What independence assumption is implied by option (b)? (*Hint*: take two different books of the same genre, are the distributions of words in these two books the same?)
6. Again with reference to option (b), consider the following query:

$$P(\text{spy}|\text{mind}, \neg\text{oven}).$$

Evaluate that query using variable elimination. According to your statistics, 80% of your books are spy novels, whereas the percentage of cookbooks in your collection is equal to the number of the month in which you were born (for example, if you were born in March, 3% of your books are cookbooks).

Bookworms



(a) (b)

7. After much tinkering with the network, you realize that a few books actually belong to two or more genres. Do you need to modify your network in order to model that fact, and if so, can you think of a suitable network?

Predictive justice?

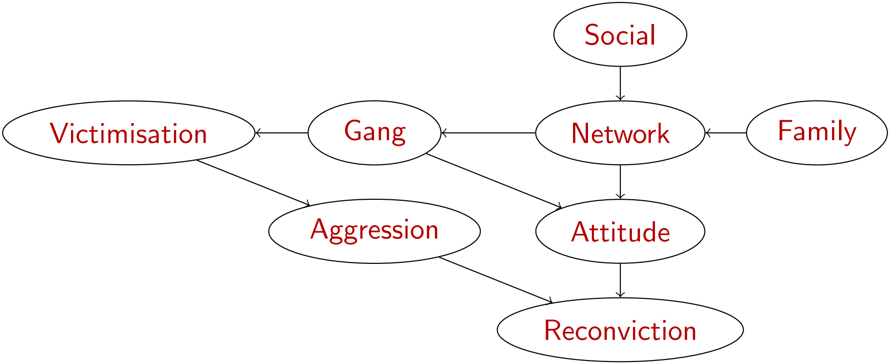
A group of forensic medics is building a Bayesian Network model to assess the risk of violent reoffending of convicted criminals, and decide possible interventions. They found out that involvement in criminal lifestyle is associated with violence, and is influenced by the presence of criminal activity in the family or peer groups, which in turn may foster attitudes supportive of crime.

Predictive justice?

Their convict model uses the following Boolean variables and connections:

- Social: presence of averse socio-economic factors
- Family: presence of a criminal family background
- Network: involvement in a criminal network
- Gang: membership in a criminal gang
- Victimisation: feeling of victimisation leading to violent thoughts
- Attitude: danger level of negative/criminal attitude
- Aggression: danger level of aggressive behaviour
- Reconviction: violent reoffence leading to reconviction

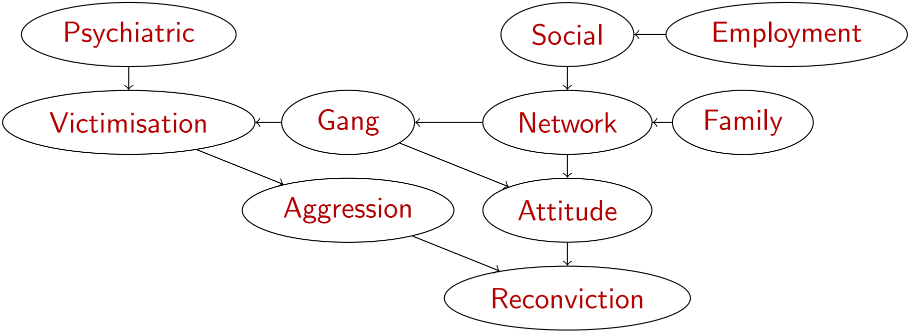
Predictive justice?



Predictive justice?

1. Define possible CPTs for Gang and Attitude. Assume it **impossible** to be a member of a Gang **without** being involved in a criminal Network.
2. According to the model, can Attitude influence Aggression of a gang member?
   Hint: a "gang member" is an individual for which Gang=true.
3. Experts define two possible interventions: Psychiatric therapy, aimed to reduce the feeling of Victimisation, and Employment policies, aimed to improve Social factors. Expand the network to include these possible interventions.
   Hint: each "intervention" can be modelled as a new Boolean variable.

Predictive justice?



4. Which interventions could reduce the risk of Reconviction of subjects known to be involved in a criminal Network?

Predictive justice?

5. Illustrate a method to compute the probability of Victimisation given (1) a dangerous level of negative/criminal attitude, (2) averse socio-economic factors, and (3) no psychiatric therapy (Attitude=true, Social=true, Psychiatric=false).

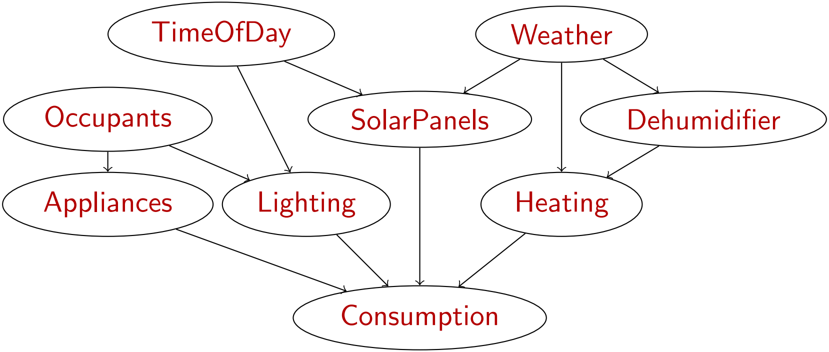
   Base your illustration on the CPTs you defined earlier, assuming P(family)=0.4, and the following CPTs:

$P(\text{network}\mid\text{social}, \text{family})$	
s, f	0.8
s, ¬f	0.5
¬s, f	0.6
¬s, ¬f	0.1

$P(\text{victimisation}\mid\text{psychiatric}, \text{gang})$	
p, g	0.4
p, ¬g	0.2
¬p, g	0.6
¬p, ¬g	0.3

Energy Efficiency

This is from the September 2024 written exam



- $\mathbf{P}(\text{Heating} \mid \text{SolarPanels}{=}\textit{high}, \text{Dehumidifiers}{=}\textit{low})$

Concluding remarks

What we didn't cover

Probabilistic Graphical Models a huge successful area.
Many things we couldn't cover

- Representation, e.g., plate models, undirected graphs
- Inference methods, e.g., belief propagation, junction tree
- Dynamic Belief Networks, e.g., HMM, Kalman filters, CRF
- Probabilistic Logic Programming
- Learning from data
- Tools and libraries
- and much more

Probabilistic Logic Programming

- Various approaches for combining LP with probability theory
- PLP enables representing complex relations among entities as well as modeling uncertainty over attributes and relations
- Promising combination of symbolic and sub-symbolic AI

Monty Hall puzzle (LPAD syntax)

```
prize(1): 1/3 ; prize(2) : 1/3 : prize(3) : 1/3.
open_door(2): 0.5 : open_door(3) : 0.5 :- prize (1).
open_door(2) :- prize (3).
open_door(3) :- prize (2).
win_keep :- prize (1).
win_switch :- prize (2), open_door(3).
win_switch :- prize (3), open_door(2).

?-prob(win_keep,Prob). % returns Prob = 0.3333333333333333
```

Software libraries

- Several libraries available (start from this list, made in 2014)
- pgmpy is one such library in Python, with basic documentation

```python
from pgmpy.models import BayesianModel
from pgmpy.factors.discrete import TabularCPD

# Defining the network structure
model = BayesianModel([('C', 'H'), ('P', 'H')])

# Defining the CPDs:
cpd_c = TabularCPD('C', 3, [[0.33, 0.33, 0.33]])
cpd_p = TabularCPD('P', 3, [[0.33, 0.33, 0.33]])
cpd_h = TabularCPD('H', 3, [[0, 0, 0, 0, 0.5, 1, 0, 1, 0.5],
                            [0.5, 0, 1, 0, 0, 0, 1, 0, 0.5],
                            [0.5, 1, 0, 1, 0.5, 0, 0, 0, 0]],
                  evidence=['C', 'P'], evidence_card=[3, 3])

# Associating the CPDs with the network structure.
model.add_cpds(cpd_c, cpd_p, cpd_h)
# Infering the posterior probability
from pgmpy.inference import VariableElimination

infer = VariableElimination(model)
posterior_p = infer.query(['P'], evidence={'C': 0, 'H': 2})
print(posterior_p['P'])
```

```
+-----+----------+
| P   |   phi(P) |
|-----+----------|
| P_0 |   0.3333 |
| P_1 |   0.6667 |
| P_2 |   0.0000 |
+-----+----------+
```

Questions?

Thank You