Fundamentals of AI and KR

Module 3: probabilistic and uncertain reasoning

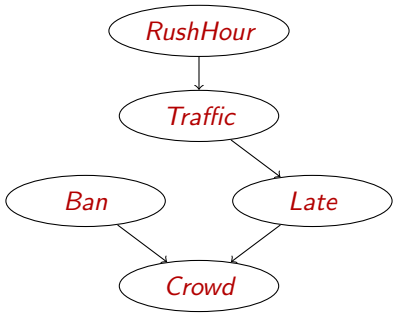
Paolo Torroni

June 16, 2020 [25 minutes for solving + 10 for uploading]

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.



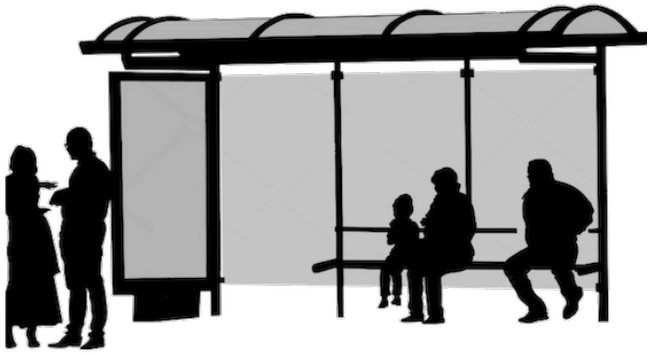
All variables are Boolean:

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

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.

1 Questions

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?
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?)
5. Is it true that P $\models$ (*Police* $\perp$ *Ban* | *Crowd*, *Deviation*, *RushHour*)?
6. What is *Deviation*'s Markov blanket?
7. Consider the following query: P(*Traffic*|*RushHour*=*False*,*Deviation*=*True*). How can you evaluate it using variable elimination? Show only the first step: P($T|\neg r,d$) $= \alpha \ldots$



2 Solution

1. Some possibly reasonable values for the CPTs:

   - P(R):

0.2	0.8

   - P(T|R):

r	**0.8**	0.2
$\neg r$	**0.4**	0.6

   - P(L|T):

t	**0.9**	0.1
$\neg t$	**0.2**	0.8

   - P(B):

2/7	5/7

   - P(C|B,L):

b	l	**0.95**	0.05
b	$\neg l$	**0.6**	0.4
$\neg b$	l	**0.5**	0.5
$\neg b$	$\neg l$	**0.1**	0.9

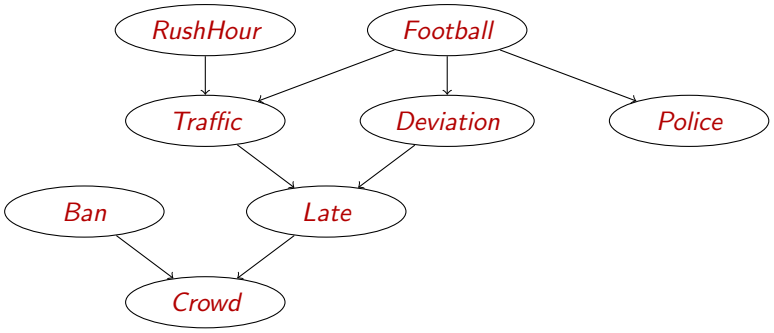
   There are $1{+}2{+}2{+}1{+}2{\times}2 = 10$ independent values (in bold)

2. By **evidential reasoning**, we seek to evaluate probable causes (RushHour? Ban?) to **explain** observations (Crowd)

   - What is the probability of RushHour given Crowd? ... P(R|C)
   - What is the probability of Ban given Crowd? ... P(B|C)

3. RushHour and Ban are not independent given Crowd. Indeed, knowing that RushHour is false increases the probability of Ban as a possible alternative for Crowd: P(b|c,¬r)>P(b|c)

4. Here is one possible way of extending the network:



5. No, because there is an active trail *Police* - *Football* - *Traffic* - *Late* - *Crowd* - *Ban*, therefore the two variables are not independent given the evidence.

6. *mb(Deviation)* = {*Football*, *Traffic*, *Late*}

7. P(*Traffic*|¬*r*,*d*) = $\alpha \sum_f$ P(*T*|*f*,¬ *r*) P(*f*) P(*d*|*f*).
   All other variables are irrelevant since they are not in the ancestry tree.