AI Assignment 3

Åsmund Brekke, Fredrik Vatsendvik

March 2017

1 Introduction

In this exercise, a decision network is made for buying used cars. The network consists of 6 hidden variables and an equal amount of observable variables. The purpose of this exercise is to model a decision network that can be used to make simple decisions in a fairly small world. The decision in the presented network is whether one should buy a particular used car or not based on the information given by described in the network.

Table 1 shows the hidden variables used an their domains. The known variables are shown in Table 2. Some assumptions are made in regards to the domains and the type of variables. For example, the variable 'Car Price' can take on values Cheap, Medium, and Expensive. The reason it is not further specified is because it allows relativeness - An agent may decide the car price based on budget, which is not modelled in this DN.

Name	Domain
Encumbrance	[True, False]
Stolen	[True, False]
Been in collision	[True, False]
Previous Owners	[Few, Many]
Functional defects	[None, Few, Many]
Structural defects	[None, Few, Many]

Table 1: Hidden variables

Name	Domain
Seller Impression	[Bad, Neutral, Good]
Mileage	[Short, Medium, Far]
Cosmetics	[Bad, Decent, Good]
Car Age	[New, Medium, Old]
Car Brand	[Volkswagen, Seat, Mercedes]
Car Price	[Cheap, Medium, Expensive]

Table 2: Observable variables

The areas selected for modelling in this exercise are related to the physical aspects of the car, such as defects, age and mileage, in addition to legal aspects such as encumbrance. The model represents Car Brand as Volkswagen, Seat and Mercedes, but these are just examples. In a 'real' model, one would have a CPT of functional defects not only of all car brands, but on individual models with regards to functional and structural defects. This is a simplification in the current model to avoid large CPT's. Another simplification is the connections between Car Brand, Age to Stolen and Encumbrance. In the real world, some car brands and models are stolen more often than others, and old cars are stolen more often than new ones [1]. We also choose to relate the seller impression with the probability of the car being a stolen car, and/or encumbered. Table 3 shows the conditional probability of the car being stolen given the impression of the seller. This seems to be the the most exaggerated CPT in the whole model, by far – but is entirely based on the use of language and the definition of a bad seller impression. If the defition of a bad seller impression is that something definitely is wrong, then a 50 percent chance of the car being stolen may seem perfectly reasonable. Another compelling argument for the authors is that this makes a bad seller impression have a larger impact on the final decision.

$P(Stolen \mid SellerImpression)$			
Seller Impression	Bad	Neutral	Good
False	0.5	0.8	0.99
True	0.5	0.2	0.01

Table 3: CPT Table for Stolen-node

The probabilities specified in the model are a result of 'guesstimation' and exaggeration to induce the intended effects in the model. For example, the CPT for $P(FunctionalDefects \mid BeenInCollision, CarBrand, CarAge)$ is given in Table 4 and Table 5, which is grossly exaggerated for effect. The real importance of this table is the difference in probabilities between the different brands of cars and the age of the car. For example, we assume that a Mercedes is a lot sturdier than a Seat, and therefore has a bigger probability of having no

[1]The National Insurance Crime Bureau 'Hot Wheels 2015' report

defects, whether both has been in a collision or not, and that the older the cars get, the greater the chances are that the cars have more defects, structural (such as rust or damages to the chassis) and functional (electrical, engine and so forth).

$P(FunctionalDefects \| BeenInCollision, CarBrand, CarAge)$									
Been In Collision	TRUE								
Car Brand	Volkswagen			Seat			Mercedes		
Car Age	New	Medium	Old	New	Medium	Old	New	Medium	Old
None	0,2	0,15	0,1	0,15	0,1	0,05	0,5	0,4	0,25
Few	0,6	0,5	0,45	0,5	0,45	0,35	0,35	0,35	0,4
Many	0,2	0,35	0,45	0,35	0,45	0,6	0,15	0,25	0,35

Table 4: CPT for $P(FunctionalDefects|BeenInCollision, CarBrand, CarAge)$ part 1

$P(FunctionalDefects \| BeenInCollision, CarBrand, CarAge)$									
Been In Collision	FALSE								
Car Brand	Volkswagen			Seat			Mercedes		
Car Age	New	Medium	Old	New	Medium	Old	New	Medium	Old
None	0,6	0,5	0,3	0,3	0,2	0,1	0,75	0,6	0,4
Few	0,3	0,4	0,5	0,6	0,6	0,5	0,2	0,3	0,4
Many	0,1	0,1	0,2	0,1	0,2	0,4	0,05	0,1	0,2

Table 5: CPT for $P(FunctionalDefects|BeenInCollision, CarBrand, CarAge)$ part 2

As the utility of the hidden states can largely be said to be separate, an aggregate of several utility functions are used in order to suggest whether the agent should buy the car or not. For example, the partial utility function for Structural Defects are given in Table 6. The separate utility functions only apply to the option of buying the car, where as the option of not buying is assigned a static value of 0.

Structural Defects	Utility value
None	35
Few	-5
Many	-45

Table 6: Utility values for Structural Defects

In Figure 1 we see an example given the observed values, where the result of the aggregated utility functions are highly positive towards the decision of buying the car. Given a good seller, and a new car in decent shape within budget.

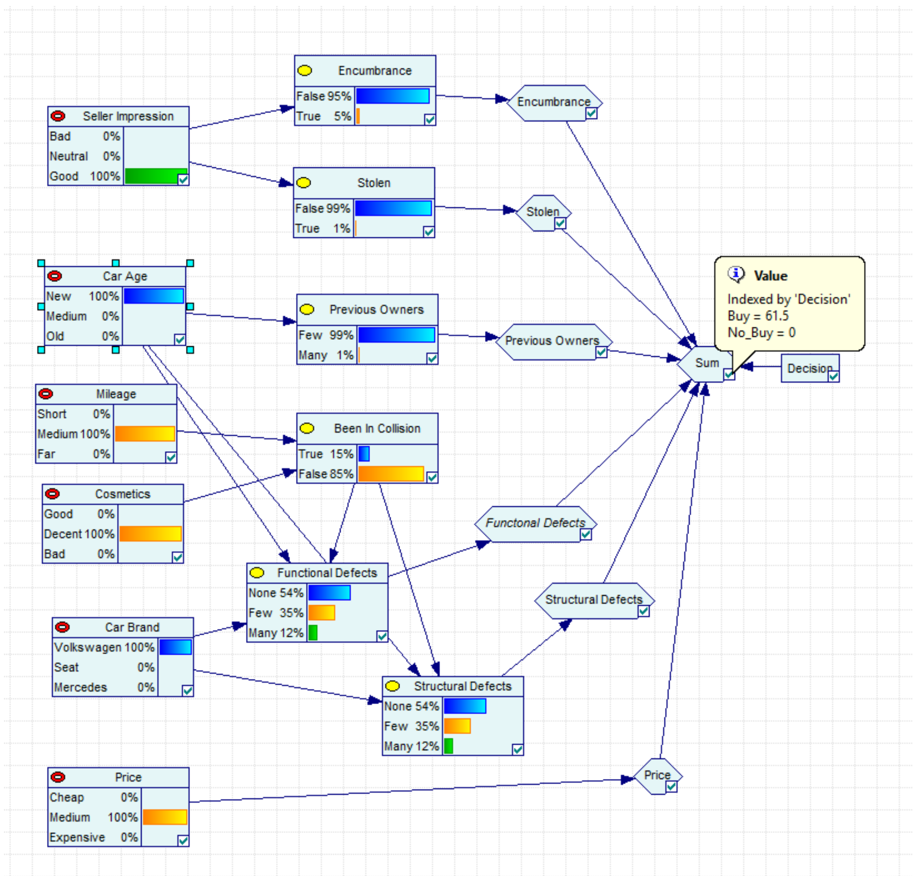

Figure 1: A case where the agent/DN chooses to buy the car.

In Figure 2 we see somewhat of an opposite scenario. Here we have the case of a bad seller, selling an old car in poor condition, of a less decirable brand. The price of the car might be cheap, but the utility functions weight this as a poor buy.

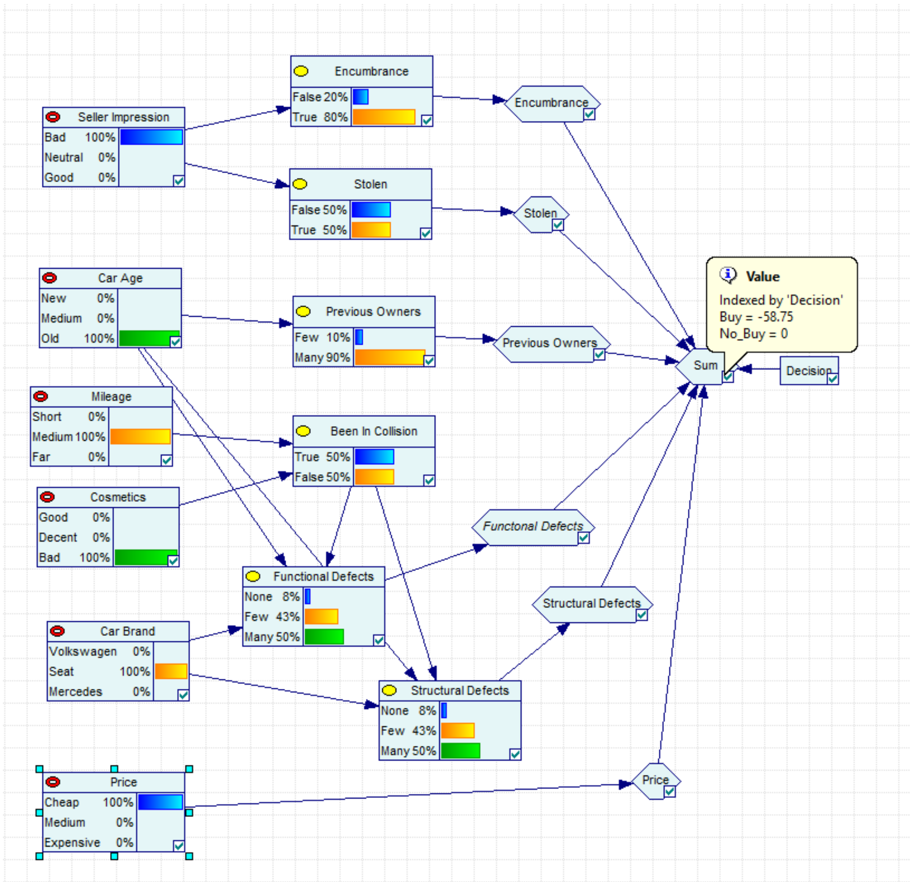

Figure 2: A case where the agent/DN chooses not to buy the car.

In Figure 3 we see a more difficult decision around a car that falls close to the mid range of options. Seller impression is neutral, and car condition is overall decent and of a desirable brand. While the car is within budget, it is old and just barely not deemed good enough to buy.

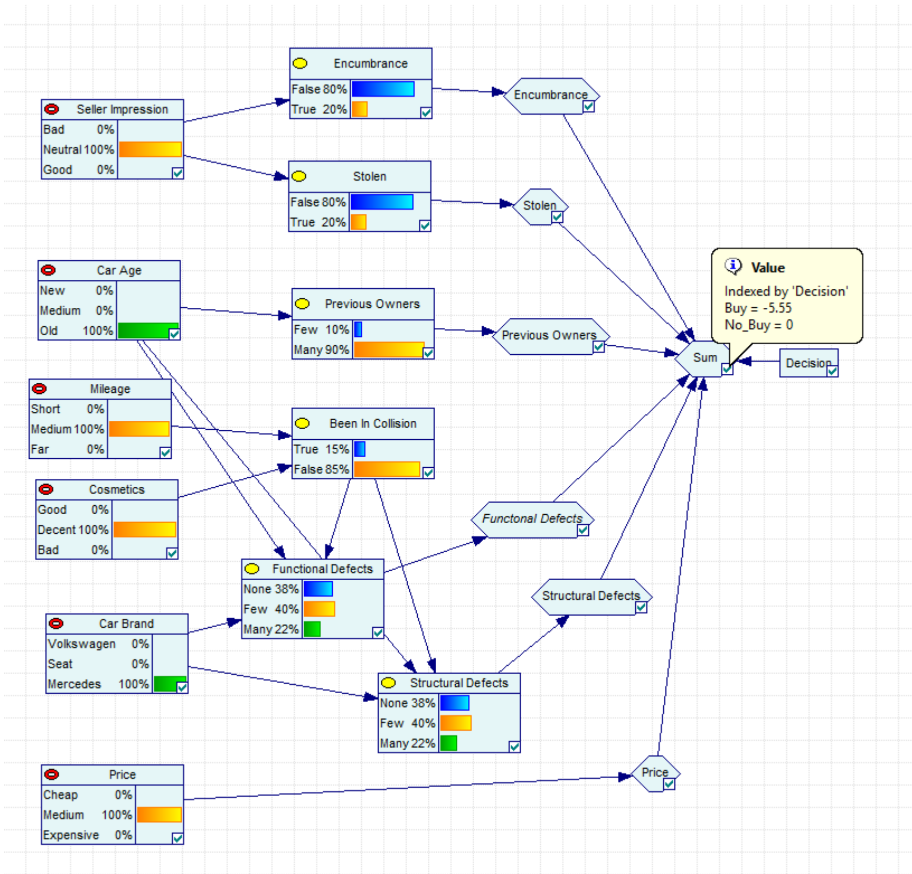

Figure 3: A case where the agent/DN just barely chooses not to buy the car.

Some ideas concerning the variables and their interconnectivedness did not make it into the final model or were changed at the very end after some discussion. One example is pertaining to the effect of a car's age on the potential for functional or structural defects. This was initially omitted to keep the model more simple, but in the end improved the reliability of the utility functions, especially around the tougher choices. For the defect tables it was proposed the addition of a medium tier, inbetween few and many, but this was scrapped due to the rather vague nature of such a quantification. Apart from this the choice of boths hidden and observed states for the model was largely agreed upon from the very start, and once the idea of determining whether to buy a used car was pitched, not many suggestions were rejected or rolled back. Apart from some minor differences in probability distributions and perhaps omission of the encumrance variable, not much would be different in this model if the members of the group undertook the task of modeling it separately.