Probabilistic Mars missions and visual representations of Donald Duck

van Hoorn, W. and Remmits, Y.
Radboud University

I. Introduction

This is the second of the two practical assignments in the series of the course Knowledge Representation and Reasoning.

For the first part of the assignment we modelled a mission to Mars, where probability can indicate whether it would be possible for microbial life to develop on Mars The second part of the assignment consists of probalistic explanations for visual information, where we model probabilistic visual features and logical axioms about the visual domain.

Our implementations are tested on Linux systems and within the Problog editor from KU Leuven website and are present in the .zip file.

II. Assignment 2-1

For the first assignment we build a knowledge system in problog based on the following story:

There have been several missions to Mars to establish, among others, whether it would be possible for microbial life to develop on Mars. Specifically, three major conditions need to be fulfilled in order for life to be possible: favourable atmospheric conditions, the presence of liquid water, and particular chemical nutrients required for microbial life. The probability of life developing if all conditions are met is estimated to be 0.01. A priori probabilities of these conditions being met are 0.1, 0.05 and 0.05 respectively, according to some expert.
The presence of chemical nutrients can be tested by performing spectrography tests, for example on soil samples taken by the Mars-rover Curiosity. These measurements are performed with a spectrometer which has a sensitivity and specficity of 90%. If liquid water has existed on Mars this would lead to particular observable geologic features with 60% probability. However, these could also have been caused independently by some other process, known to occur on Mars. This process is estimated to cause the geological features with a probability of 20%. If microbial life were present there would be a 90% chance that this would produce elevated levels of methane, which would otherwise only have 3% chance to be observed. Methane can also be measured using a spectrometer, this one with a sensitivity of 97% and specificity of 92%.

The corresponding baysian network can be found in appendix I. The following probability tables correspond to this network

Nutri	P(Nutri)
true	0.05

Atmos	P(Atmos)
true	0.1

Water	P(Water)
true	0.05

Process	P(Process)
true	1

Nutri	P(Specnut | Nutri)
true	0.9
false	0.1

Water	Process	P(Geo | Water, process)
true	true	0.68
false	true	0.2
true	false	0.6
false	false	0

Nutri	Atmos	Water	P(Life | Nurti, Water, Atmos)
true	true	true	0.01
true	true	false	0
true	false	true	0
true	false	false	0
false	true	true	0
false	true	false	0
false	false	true	0
false	false	false	0

Life	P(Meth | Life)
true	0.9
false	0.03

Meth	P(Specmet | Meth)
true	0.97
false	0.08

The Problog code that is corresponding to the network and the probability tables can be found in appendix II. Observations influence the probability of microbial life on Mars. For example evidence of for any single feature improves the probability of life on Mars. When there is evidence of Nutri, Water or Meth, evidence of respectively Specnut, Geo and Specmet becomes irrelevant to the probability of life, since these factors are conditionally independent of life.

Problog has an implicit calculation mechanism which it uses as the explain capability. Problog calculates the probability of, for example, Geo by calculating a noisy-OR over the apriori probabilities for Process and Water and combining these. You could also calculate the probability for Geo with baysian probability theory.

Using Problog to represent probabilistic knowledge has the main advantage that this representation provides for efficient computation of probabilities. In addition it is very easy to add nodes and/or evidence to the model, without changing much to the previous model. On the other hand the Problog representation is less insightfull. The visual aspect of a baysian network makes it very easy to understand the model and (conditional) independencies.

The decribed model only offers the apriori probabilities of one expert, while it would be likely that different experts have different opinions on this. This can be solved by using the first-order logic probabilities of Problog. This addaption of the model can be fount in appendix III. The probabilities of multiple experts can now simultataneously affect the probability of life on Mars. Since two experts now add evidence to nutrients, water and atmosphere for being present, these multiple expert increase the probability of life on Mars. We can represent these multiple experts in the bayian network as causal nodes to Nutri, Atmos and Water.

A second addition to the model might be the possibility of adding multiple observations, for example multiple measures with a spectrometer. This can also be solved with the first-order capabilities of Problog. We can add multiple evidence for spectrometer readings. Problog combines this combinated evidence and this influences the probability of life on Mars. We can represent these multiple measurements in the baysian network as multiple spectrometer nodes under Nutri and Meth.

To model dynamic conditions and their effect on life we used the suggested simplified model where there are three conditions (liquid water, nutrients and specific atmospheric conditions) needed for life to exist on Mars. We assume life and these three conditions to be present at t_0. Then the probability of survival at time t_n given the previous situatio with a number n of satisfied conditions, is given by the following formula:

$$P(life(t)|life(t-1), conditions(t-1) = n) = \begin{cases} 1 - \frac{1}{n+1} & n > 0 \\ 0 & otherwise \end{cases} \tag{1}$$

The probability that the number of present conditions at t remains equal to the number of conditions at $t-1$ is 0.8, while the probability of this mumber increasing or decreasing by one is 0.1 for both.

To model this we used the following predicates Life(T) and Conditions(T, N) with the initial condition Life(0) and Conditions(0, 3). In order to implement this we need to consider the boundary situations where the amount of satisfied conditions is 3, where is can not become higher, or 0 where it can not become lower. In these boundary situations, we added the probability of n increase or respectively decrease to the probability of n remaining the same. This results in a transition distribution of 0.9 for n to remain n and 0.1 for n to decrease or increase depending on the specific boundary situation. See code attached

Given that t = 0 is the initial situation, the probability of life at $t = 1$ is 0.75, the probability of life at $t = 2$ is 0.556 and the probability of life at $t = 5$ is 0.215.

We can see that the probability of life on mars is decreasing drastically over time, implying little chance of survival of life on mars.

III. Assignment 2-2

As a tribute to our childhood we chose to specify the domain of Donald Duck characters. Donald Duck is a Disney character who became so popular that he has an universe of his own. This domain contains all citizens of Ducksburg, but since this are too many characters for the scope of the assignment we limited the domain to, what we believe are, the eight most popular characters of Donald Duck characters. These characters are Donald Duck, Daisy Duck, Scrooge McDuck, the three nephews (Huey (kwik), Dewey (kwek) and Louie (kwak) Duck), Glandstone Galder and Elviry Coot aka Grandma Duck.

The specified characters are all ducks thus we do not need to represent animal types. Therefore we rely on clothing properties and character size. To make it more interesting for Problog we actively chose characters with overlapping properties between subjects and varying properties within subjects.

Donald usually wears a blue shirt, a bow tie and a blue hat. Other illustrators gave donald a black shirt instead. Also Donald has worn many uniform from all his different jobs thus we also added an apriori probablitiy for any color of shirt. Daisy usually dresses in pink with a pink ribbon in her hair and as far as we have seen Daisy, she always wears shoes. The nephews already look alike because their only difference is the color of their shirt and hat. We added the length to the representation since the nephews are smaller than anybody else. We also added a small probability for Daisy and Donald to be small since there are some comics in which they are baby's. Scrooge McDuck, Gladstone Galder and Grandma Duck wear spats besides the fact that they wear outfits with different colors. Also Scrooge and Grandma wear glasses, Scrooge and Gladstone wear a hat and Grandma wears shoes under her spats. These are the most common appearances of these characters, however illustrators can let them wear whatever they want thus we added apriori probabilities for all of the properties.

We choose the probabilities is such way that the known features for each duck are not true which a probability of one, because there is always a chance for every feature to be different or not present in the picture.

The style of reasoning we employed in the representation is abductive of nature. As mentioned in the assignment, deductive reasoning raises a problem when prior probablities are defined. That is, when more evidence is availible it becomes less likely to be a particular cause. This worried us as we forsaw that we might need prior probabilities in the domain of Donald Duck characters, thus we wanted to be prepared and employ abductive reasoning.

As we identify an image we query all characters given the observed evidence. We chose to give Donald Duck a slightly higher apriori probability because we believe the main character will occur more in regard of the other characters. When we, for instance, observe a blue shirt we find the highest probability for Donald, with a probability of 0.452, and Dewey (kwek) comes second with a probability of 0.339. Scrooge McDuck has a probability of 0.101 while the other characters, who only hold the apriori probability of wearing a blue shirt, have a probability of 0.021.

It is more interesting to find shifts. When we query with spats observed we find that it is equally likely that look at Grandma or at Gladstone with a probabiltity of .337, Scrooge with 0.307 and the others > 0.005. However, when we also observe a red shirt we see the probability of Grandma drop to a mere 0.031. Gladstone has now the highest probability, which is 0.503 and Scrooge is second with 0.458. The others have a probability of > 0.001 except for Huey (kwik) because he wears a red shirt too. Huey has a probabilty of 0.005.

We wanted to test our representation with a difficult image, as showed in appendix IX figure 11. The character in this image is difficult to identify because Scrooge is not wearing his usual outfit. On this image Scrooge is wearing a pink and black striped bathing suit without a hat or spats. When we queried on this image with a black shirt, a pink shirt and glasses as evidence, we did not receive the desired result, as expected. Scrooge received a probability of 0.05 and Grandma received a probability of 0.967.

This result shows us that our representation is not sufficient enough to identify Scrooge McDuck and thus will not identify all characters in a sufficient manner. We might want to consider to add more features to the model such as hairstyles, gender or external features, for instance the presence of money in case of Scrooge McDuck.

We tried adding probability in the evidence, since it would make sense that computer vision is never fully sure that is observes glasses from two circles. However we were unable to implement this feature.

Extending the model with more characters is quite easily done. However making it more realistic is much harder. To make it more realistic we would have to involve more than the higher level features we use now. In order for this model to be used in automatic recognition of the Ducks, firstly the algorithm needs to be able to model lower order features like shapes, one layer above that it would have to be able to model the shapes of shirts hats and glasses. When al these features are incorporated in the model, we think a model like this would be able to predict the identity of the characters correctly in such a small domain as the Duck world.

IV. Reflection

I. Assignment 2-1

For the first assignment it took us approximately twenty hours each to finish the assignment, this was due to a lot time dedicated to bug fixing with assistence. What we missed was clear idea of the Problog syntax other than the tutorial examples. However we must admit that we did not spent a lot of time on the availible tutorials. We found Problog ideal for the representation of Bayesian networks and intend to use Problog again when we need to build another Bayesian network again.

II. Assignment 2-2

At the moment of writing we are ten hours into the second assignment and we need a couple of hours more to wrap it all up. At first we interpreted the assignment wrong, by our mistake. We thought we needed to be able to read images with Problog but that seemed us impossible within the time scope. Luckily we found out what we actually needed to do before we implemented anything. We think the goals of the second assignment might have been clearer. It is unclear when the subject is good enough or detailed well enough.

V. Appendix

I.

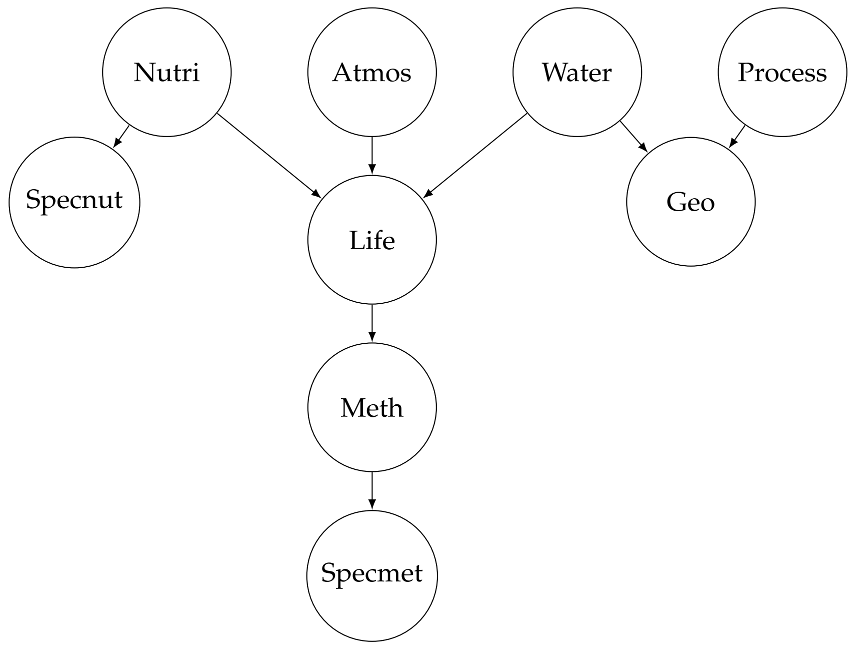

Figure 1: *Baysian Network for Probabilistic Mars mission*

II.

```
%Mars mission query geological features,
%assuming no evidence

%A priori
1.00::process.
0.05::water.
0.10::atmos.
0.05::nutri.

%Spectrometer for nutrition
0.90::specnut :- nutri.
0.10::specnut :- \+nutri.

0.68::geo :- process, water.
0.20::geo :- process, \+water.
0.60::geo :- \+process, water.

0.01::life :- nutri, water, atmos.

0.90::meth :- life.
0.03::meth :- \+life.

%Spectrometer for methane
0.97::specmeth :- meth.
0.08::specmeth :- \+meth.

query(geo).
```

III.

```
%Mars mission query microbial life,
%allowing multiple experts and
%evidence of multiple spectrography measurements

%A priori
1.00::process.

expert(0.05, 0.10, 0.05).
expert(0.15, 0.20, 0.15).

X::water :- expert(X,Y,Z).
Y::atmos :- expert(X,Y,Z).
Z::nutri :- expert(X,Y,Z).

%Spectrometer for nutrition
0.90::specnut(M) :- nutri.
0.10::specnut(M) :- \+nutri.

0.68::geo :- process, water.
0.20::geo :- process, \+water.
0.60::geo :- \+process, water.

0.01::life :- nutri, water, atmos.

0.90::meth :- life.
0.03::meth :- \+life.

%Spectrometer for methane
0.97::specmeth :- meth.
0.08::specmeth :- \+meth.

evidence(specnut(1), true).
evidence(specnut(2), true).
evidence(specnut(3), false).

query(life).
```

IV.

```
%Dynamic Mars mission queries microbial life(S) for S=1,2,5

%A priori
life(0).
conditions(0, 3).

%Calculates P of life(S) given life(Smin)
%and conditions(Smin, N)
calcP(N, P):- N>0, P is 1-(1/(N+1)).

%lower bound of n
bound0(0, 0).
bound0(N, Nnew) :- N>0, Nnew is N-1.

%upper bound of n
bound3(3, 3).
bound3(N, Nnew) :- N<3, Nnew is N+1.

%Calculates N
0.80::conditions(S, N);
0.10::conditions(S, Npl);
0.10::conditions(S, Nmin) :-
S>0,
Smin is S-1,
conditions(Smin, N),
bound0(N, Nmin),
bound3(N, Npl).

P::life(S) :- S>0, Smin is S-1, life(Smin), conditions(Smin, N), calcP(N, P).

query(life(1)).
query(life(2)).
query(life(5)).
```

V.

```
%Donald Duck domein

%A priori
0.16::donald(X) ;
0.12::dagobert(X) ;
0.12::katrien(X);
0.12::kwik(X);
0.12::kwek(X);
0.12::kwak(X);
0.12::guus(X);
0.12::oma(X).

%A priori
0.05::shirt(blauw, X).
0.05::shirt(paars, X).
0.05::shirt(roze, X).
0.05::shirt(rood, X).
0.05::shirt(oranje, X).
0.05::shirt(geel, X).
0.05::shirt(groen, X).
0.05::shirt(wit, X).
0.05::shirt(zwart, X).

%A priori
0.05::hoed(blauw, X).
0.05::hoed(paars, X).
0.05::hoed(roze, X).
0.05::hoed(rood, X).
0.05::hoed(oranje, X).
0.05::hoed(geel, X).
0.05::hoed(groen, X).
0.05::hoed(wit, X).
0.05::hoed(zwart, X).

%A priori
0.01::schoen(X).
0.01::slobkous(X).
0.01::strik(X).
0.01::klein(X).
0.01::bril(X).

% Kansen van characters gebaseerd op schatting n.a.v. image search op Google
% Nooit een kans van 1.0 omdat dit niet realistisch is

%Donald
0.90::hoed(blauw, X) :- donald(X).
```

```
0.99::strik(X) :- donald(X).
0.79::shirt(blauw, X); 0.20::shirt(zwart, X) :- donald(X).
0.05::klein(X) :- donald(X).

%Dagobert
0.80::hoed(zwart, X):- dagobert(X).
0.99::bril(X) :- dagobert(X).
0.79::shirt(rood, X); 0.20::shirt(blauw, X) :- dagobert(X).
0.90::slobkous(X) :- dagobert(X).

%Katrien
0.99::hoed(roze, X) :- katrien(X).
0.99::strik(X) :- katrien(X).
0.79::shirt(roze, X); 0.20::shirt(zwart, X) :-katrien(X).
0.99::shoen(X) :- katrien(X).
0.05::klein(X) :- katrien(X).

%Kwik
0.80::hoed(rood, X) :- kwik(X).
0.79::shirt(rood, X); 0.20::shirt(zwart, X) :- kwik(X).
0.99::klein(X) :-kwik(X).

%Kwek
0.80::hoed(blauw, X) :- kwek(X).
0.79::shirt(blauw, X); 0.20::shirt(zwart, X) :- kwek(X).
0.99::klein(X) :-kwek(X).

%Kwak
0.80::hoed(groen, X) :- kwak(X).
0.80::shirt(groen, X); 0.20::shirt(zwart, X) :- kwak(X).
0.99::klein(X) :-kwak(X).

%Guus Geluk
0.20::hoed(rood, X); 0.70::hoed(groen, X) :- guus(X).
0.99::strik(X) :- guus(X).
0.79::shirt(rood, X); 0.20::shirt(groen, X) :- guus(X).
0.99::slobkous(X) :- guus(X).
0.20::klaver(X) :- guus(X).

%Oma Duck
0.99::bril(X) :- oma(X).
0.89::shirt(zwart, X); 0.10::shirt(paars, X) :- oma(X).
0.99::slobkous(X) :- oma(X).
0.99::schoen(X) :- oma(X).

evidence(shirt(blauw, a)).

query(donald(a)).
```

```
query(dagobert(a)).
query(katrien(a)).
query(kwik(a)).
query(kwek(a)).
query(kwak(a)).
query(guus(a)).
query(oma(a)).
```

VI.

Figure 2: *Donald Duck*

Figure 3: *Baby Donald*

VII.

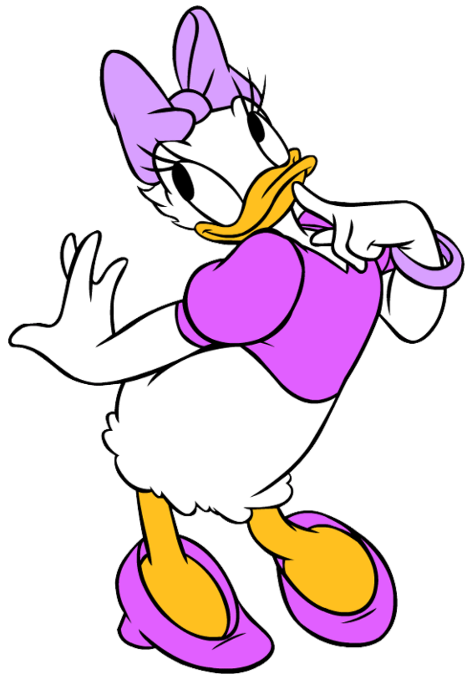

Figure 4: *Daisy Duck*

Figure 5: *Baby Daisy*

VIII.

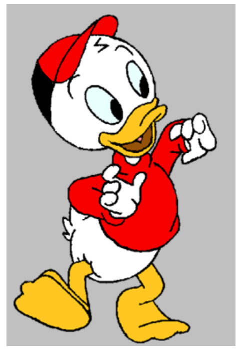

Figure 6: *Huey Duck*

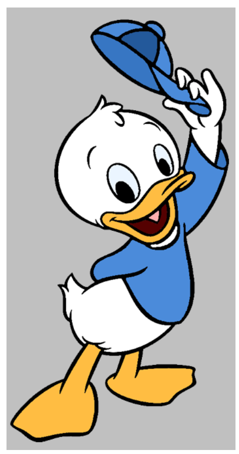

Figure 7: *Dewey Duck*

Figure 8: *Louie Duck*

IX.

Figure 9: *Scrooge McDuck in blue*

Figure 10: *Scrooge McDuck in red*

Figure 11: *Scrooge McDuck in bathing suit*

X.

Figure 12: *Gladstone in green*

Figure 13: *Gladstone in red*

XI.

Figure 14: *Grandma in purple*

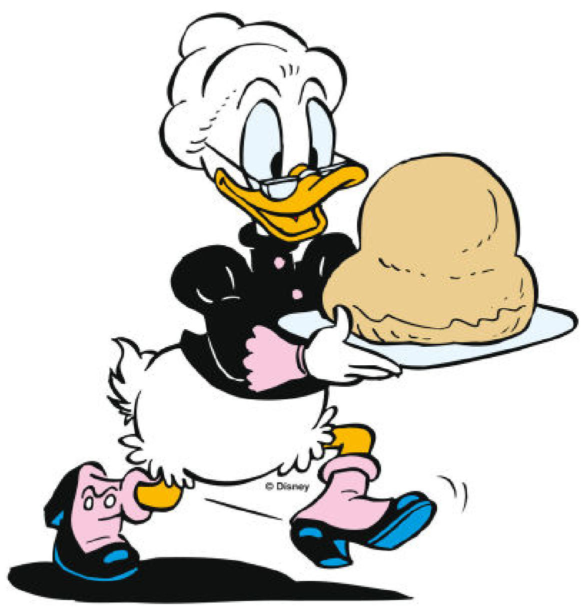

Figure 15: *Grandma in black*

XII.

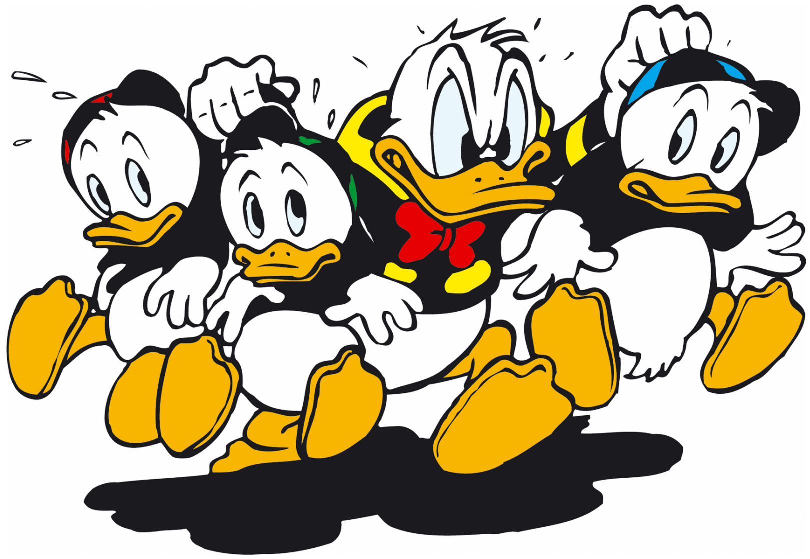

Figure 16: *Donald and nephews in black*