ARTIFICIAL INTELLIGENCE (E016350A)
GHENT UNIVERSITY
AY 2024/2025
Aleksandra Pizurica
Asst: Nicolas Nadisic
E-mail: ai@lists.ugent.be

Solutions: Rational decisions

1. The Surprise Candy Company makes candy in two flavors: 70% are strawberry flavor and 30% are anchovy flavor. Each new piece of candy starts out with a round shape; as it moves along the production line, a machine randomly selects a certain percentage to be trimmed into a square; then, each piece is wrapped in a wrapper whose color is chosen randomly to be red or brown. 80% of the strawberry candies are round and 80% have a red wrapper, while 90% of the anchovy candies are square and 90% have a brown wrapper. All candies are sold individually in sealed, identical, black boxes.

   Now you, the customer, have just bought a Surprise candy at the store but have not yet opened the box. Consider the three Bayes nets in figure below.

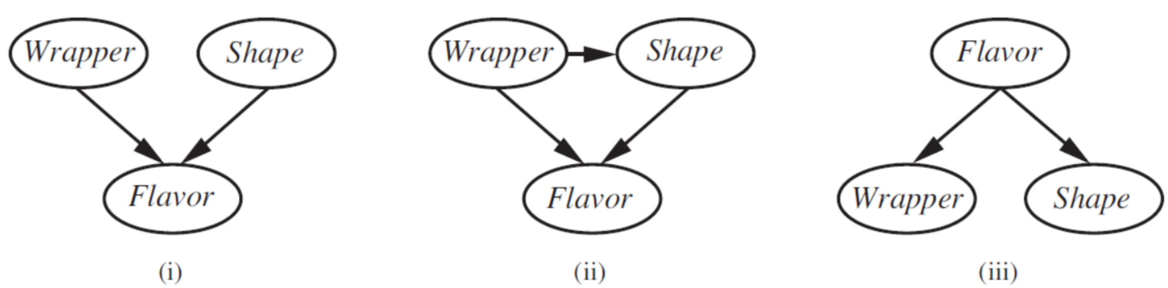


Figure 1: Three proposed Bayes nets for the Surprise Candy problem.

   (a) Which network(s) can correctly represent $\mathbf{P}(Flavor, Wrapper, Shape)$?

   (b) Which network is the best representation for this problem?

   (c) Does network (i) assert that $\mathbf{P}(Wrapper|Shape) = \mathbf{P}(Wrapper)$?

   (d) What is the probability that your candy has a red wrapper?

   (e) In the box is a round candy with a red wrapper. What is the probability that its flavor is strawberry?

   (f) A unwrapped strawberry candy is worth s on the open market and an unwrapped anchovy candy is worth a . Write an expression for the value of an unopened candy box.

   (g) A new law prohibits trading of unwrapped candies, but it is still legal to trade wrapped candies (out of the box). Is an unopened candy box now worth more than, less than, or the same as before?

Solution:

(a) Networks (ii) and (iii) can represent this network but not (i).

(ii) is fully connected, so it can represent any joint distribution.

(iii) follows the generative story given in the problem: the flavor is determined (presumably) by which machine the candy is made by, then the shape is randomly cut, and the wrapper randomly chosen, the latter choice independently of the former.

(i) cannot represent this, as this network implies that the wrapper color and shape are marginally independent, which is not so: a round candy is likely to be strawberry, which is in turn likely to be wrapped in red, whilst conversely a square candy is likely to be anchovy which is likely to be wrapped in brown.

(b) (iii) is the best. Unlike (ii), (iii) has no cycles which we have seen simplifies inference. Its edges also follow the causal direction (from “causes” to “symptoms”) so probabilities will be easier to elicit. Indeed, the problem statement has already given them.

(c) Yes, because $Wrapper$ and $Shape$ are d-separated (check *L7.2. Graphical Models And Inference*). In general when discussing problems of absolute and conditional independence, all the reasoning should be based on network separation theory.

(d) Once we know the $Flavor$ we know the probability its wrapper will be red or brown. So we marginalize $Flavor$ out:

$$\begin{aligned} P(Wrapper = red) &= \sum_f P(Wrapper = red, Flavor = f) \\ &= \sum_f P(Flavor = f)P(Wrapper = red \mid Flavor = f) \\ &= 0.7 \times 0.8 + 0.3 \times 0.1 = 0.59. \end{aligned}$$

(e) We apply Bayes theorem, by first computing the joint probabilities:

$$\begin{aligned} &P(Flavor = strawberry, Shape = round, Wrapper = red) \\ &= P(Flavor = strawberry)P(Shape = round \mid Flavor = strawberry) \\ &\quad P(Wrapper = red \mid Flavor = strawberry) \\ &= 0.7 \times 0.8 \times 0.8 = 0.448, \end{aligned}$$

$$\begin{aligned} &P(Flavor = anchovy, Shape = round, Wrapper = red) \\ &= P(Flavor = anchovy)P(Shape = round \mid Flavor = anchovy) \\ &\quad P(Wrapper = red \mid Flavor = anchovy) \\ &= 0.3 \times 0.1 \times 0.1 = 0.003. \end{aligned}$$

Normalizing these probabilities yields that it is strawberry with probability $0.448/(0.448 + 0.003) \approx 0.9933$.

Solution:
(Continued)

(f) Its value is the probability that you have a strawberry upon unwrapping times the value of a strawberry, plus the probability that you have a anchovy upon unwrapping times the value of an anchovy or

$$0.7s + 0.3a.$$

(g) The value is the same, by the axiom of decomposability (*Axiom of decomposability:* An agent is indifferent between lotteries that have the same probabilities over the same outcomes, even if one or both is a lottery over lotteries.).

2. In 1713, Nicolas Bernoulli stated a puzzle, now called the St. Petersburg paradox, which works as follows. You have the opportunity to play a game in which a fair coin is tossed repeatedly until it comes up heads. If the first heads appears on the n-th toss, you win \$$2^n$.

   (a) Show that the expected monetary value of this game is infinite.

   (b) How much would you, personally, pay to play the game?

   (c) Nicolas's cousin Daniel Bernoulli resolved the paradox in 1738 by finding that the utility of money is logarithmic (i.e., $U(S_n) = a \log_2 n + b$ where S_n is the state of having \$$n$). Assume $k = c$ for simplicity, i.e., we paid all we have to play the game. What is the expected utility of the game under this assumption?

   (d) What is the maximum amount it would be rational to pay to play the game, assuming that one's initial wealth is \$$k$?

Solution:

(a) First, the expected monetary value or EMV in statistics is the average weighted probability of number of outcomes. This gives an average of the best and worse case scenario. To answer this question, first note that the probability that the first head appears on the n-th toss is 2^{-n}. Then in order to obtain EMV, one needs to consider what would be the average payout: with probability $1/2$ the player wins 2 dollars, with probability $1/4$ the player wins 4 dollars, ... and with probability $1/2^n$ the player wins 2^n dollars, so

$$EMV(L) = \sum_{n=1}^{\infty} \frac{1}{2^n} \cdot 2^n = \sum_{n=1}^{\infty} 1 = \infty.$$

Based only on this answer, one should therefore play the game at any price if offered the opportunity.

(b) Typical answers of people asked range between \$4 and \$100. The paradox lies here, in the discrepancy between what people seem willing to pay to enter the game and the infinite expected value.

(c) The paradox has been solved by Bernouli since he said that the determination of the value of an item must not be based on the price, but rather on the utility it yields. He suggested that the utility of money is logarithmic. So, let us assume that our initial wealth is k and after paying c to play the game, we are left with $\$(k-c)$. Then

$$U(L) = \sum_{n=1}^{\infty} 2^{-n} \cdot (a \log_2 (k - c + 2^n) + b).$$

Assume $k - c = \$0$ for simplicity, i.e., we paid all we have to play the game. Then

$$\begin{aligned} U(L) &= \sum_{n=1}^{\infty} 2^{-n} \cdot (a \log_2 (2^n) + b) \\ &= \sum_{n=1}^{\infty} \frac{an}{2^n} + b \sum_{n=1}^{\infty} \frac{1}{2^n} = a \underbrace{\left(\sum_{n=1}^{\infty} \frac{n}{2^n} \right)}_{:=A} + b \underbrace{\left(\sum_{n=1}^{\infty} \frac{1}{2^n} \right)}_{:=B} \\ &= 2a + b. \end{aligned}$$

Now let us show how we obtained that $A = 2$ and $B = 1$. Note that we use the formula for the geometric series, i.e., we know that

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \qquad |x| < 1.$$

Solution:
(Continued) Then we easily obtain

$$B = \sum_{n=1}^{\infty} \frac{1}{2^n} = \sum_{n=0}^{\infty} \frac{1}{2^n} - 1 = \frac{1}{1-\frac{1}{2}} - 1 = 2 - 1 = 1$$

and then

$$\frac{1}{2}A = A - \frac{1}{2}A = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \cdots = \frac{1}{2}\left(1 + \frac{1}{2} + \frac{1}{2^2} + \dots\right) = \frac{1}{2} \cdot 2 = 1$$

and thus we conclude that $A = 2$.

(d) In order to obtain what would be rational to pay we need to check when is the utility of wealth before the event equal to the utility of wealth after the event. The maximum amount c is thus given as the solution of

$$a \log_2 k + b = \sum_{n=1}^{\infty} 2^{-n}(a \log_2 (k - c + 2^n) + b).$$

In our simple case where we assumed that $k = c$, we get

$$a \log_2 c + b = 2a + b,$$

or after solving this equation that $c = \$4$.

3. Consider a student who has the choice to buy or not buy a textbook for a course. We'll model this as a decision problem with one Boolean decision node, B, indicating whether the agent chooses to buy the book, and two Boolean chance nodes, M, indicating whether the student has mastered the material in the book, and P, indicating whether the student passes the course. Of course, there is also a utility node, U. A certain student, Sam, has an additive utility function: 0 for not buying the book and $-\$100$ for buying it; and \$2000 for passing the course and 0 for not passing. Sam's conditional probability estimates are as follows:

- $\mathrm{P}(p \mid b, m) = 0.9$
- $\mathrm{P}(p \mid b, \neg m) = 0.5$
- $\mathrm{P}(p \mid \neg b, m) = 0.8$
- $\mathrm{P}(m \mid b) = 0.9$
- $\mathrm{P}(m \mid \neg b) = 0.7$
- $\mathrm{P}(p \mid \neg b, \neg m) = 0.3$

You might think that P would be independent of B given M, but this course has an openbook final-so having the book helps.

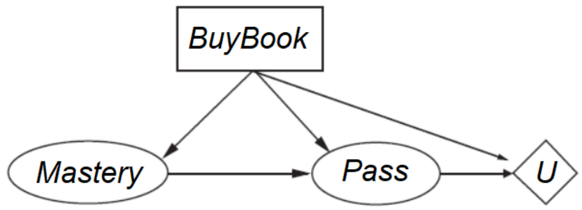
(a) Draw the decision network for this problem.

(b) Compute the expected utility of buying the book and of not buying it.

(c) What should Sam do?

Solution:

(a) A decision network for the book-buying problem is given by:



(b) First we need to calculate the posterior probabilities of all chance node parents of the utility node into which the action node feeds. We have here that the only chance node that is a parent of the utility node is *Pass*. For each of $B = b$ and $B = \neg b$, we compute $\mathrm{P}(p \mid B)$ and thus $\mathrm{P}(\neg p \mid B)$ by marginalizing out M, then use this to compute the expected utility. First:

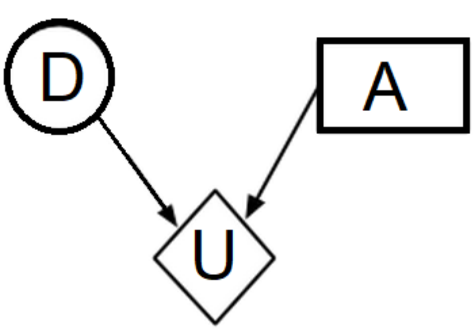
$$\begin{aligned}
\mathrm{P}(p \mid b) &= \sum_m \mathrm{P}(p \mid b, m)\mathrm{P}(m \mid b) = 0.9 \cdot 0.9 + 0.5 \cdot 0.1 = 0.86, \\
\mathrm{P}(p \mid \neg b) &= \sum_m \mathrm{P}(p \mid \neg b, m)\mathrm{P}(m \mid \neg b) = 0.8 \cdot 0.7 + 0.3 \cdot 0.3 = 0.65.
\end{aligned}$$

The expected utilities are thus:

$$\begin{aligned}
EU[b] &= \sum_p \mathrm{P}(p \mid b)U(p, b) = 0.86(2000 - 100) + 0.14(-100) = 1620, \\
EU[\neg b] &= \sum_p \mathrm{P}(p \mid \neg b)U(p, \neg b) = 0.65 \cdot 2000 + 0.14 \cdot 0 = 1300.
\end{aligned}$$

(c) Buy the book, Sam.

4. **(Old exam question)** It is Sunday evening and Sam is finishing up preparing for the AI exam that is coming up on Monday. Sam has already mastered all the topics except one: Decision Networks. He is contemplating whether to spend the remainder of his evening learning that topic ($A = learn$), or just go to sleep ($A = sleep$). Decision Networks are going to be on the exam ($D = d$) or not ($D = \neg d$) with equal *a-priori* probabilities. The student's utility of satisfaction is only affected by these two random variables D and A, as shown below:



D	$\mathbf{P}(D)$
d	0.5
$\neg d$	0.5

D	A	$U(D,A)$
d	$learn$	1400
$\neg d$	$learn$	800
d	$sleep$	0
$\neg d$	$sleep$	2000

(a) Compute the following, where EU is expected utility and MEU is maximum expected utility:

(i) $EU(learn) =$

(ii) $EU(sleep) =$

(iii) $MEU(\{\}) =$

(iv) Action that achieves $MEU(\{\}) =$

(b) The teaching assistant's happiness (H) on the same day is affected by whether decision networks are going to be on the exam. The happiness (H) influences the chance that the assistant posts on Facebook (f) or doesn't post on Facebook ($\neg f$): if happy, he posts with probability 0.6 and posts with probability 0.2 if unhappy. The prior on D and utility tables remain unchanged. Assume the prior probabilities $\mathrm{P}(h) = 0.6$ and $\mathrm{P}(f) = 0.44$. Further on, the assistant is happy with probability 0.95 if the decision networks are on the exam and is happy with probability 0.25 if they are not on the exam. Some tables that are computed from the data above and that might be useful in answering the questions below are:

D	F	$\mathbf{P}(D \mid F)$
d	f	0.666
$\neg d$	f	0.334
d	$\neg f$	0.370
$\neg d$	$\neg f$	0.630

F	D	$\mathbf{P}(F \mid D)$
f	d	0.586
$\neg f$	d	0.414
f	$\neg d$	0.300
$\neg f$	$\neg d$	0.700

(i) Draw the decision network that reflects this extended description of the problem.

(ii) Compute the expected utilities of the two actions in cases when the assistant posts or not on Facebook:

- $EU(learn \mid f) =$
- $EU(sleep \mid f) =$
- $EU(learn \mid \neg f) =$
- $EU(sleep \mid \neg f) =$

(iii) What are the maximum expected utilities and what are the corresponding optimal actions?

- $MEU(\{f\}) =$
- Optimal Action$(\{f\}) =$
- $MEU(\{\neg f\}) =$
- Optimal Action$(\{\neg f\}) =$

Solution:

(a) We can easily obtain the following values:

(i) By using the definitions (see *L.10.1. Rational Decisions*) we obtain:

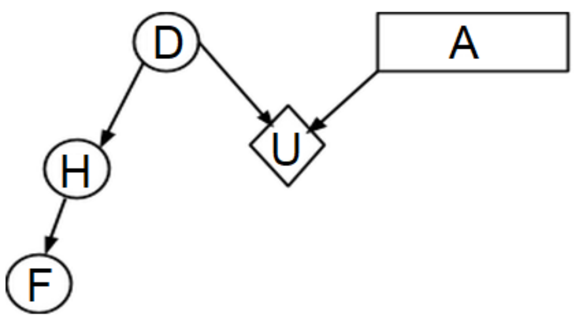
$$\begin{aligned} EU(learn) &= \sum_d \mathrm{P}(d)U(learn, d) \\ &= \mathrm{P}(d)U(d, learn) + \mathrm{P}(\neg d)U(\neg d, learn) \\ &= 0.5 \cdot 1400 + 0.5 \cdot 800 = 1100 \end{aligned}$$

(ii) $EU(sleep) = \mathrm{P}(d)U(d, sleep) + \mathrm{P}(\neg d)U(\neg d, sleep) = 0.5 \cdot 0 + 0.5 \cdot 2000 = 1000$

(iii) $MEU(\{\}) = \max(EU(learn), EU(sleep)) = \max(1100, 1000) = 1100$

(iv) Action that achieves $MEU(\{\}) = learn$.

(b) (i) From the information provided in the exercise, we can obtain the following decision network and CPT's:



D	$\mathbf{P}(D)$
d	0.5
$\neg d$	0.5

D	A	$U(D, A)$
d	$learn$	1400
$\neg d$	$learn$	800
d	$sleep$	0
$\neg d$	$sleep$	2000

H	$\mathbf{P}(H)$
h	0.6
$\neg h$	0.4

F	$\mathbf{P}(F)$
f	0.44
$\neg f$	0.56

F	H	$\mathbf{P}(F \mid H)$
f	h	0.6
$\neg f$	h	0.4
f	$\neg h$	0.2
$\neg f$	$\neg h$	0.8

H	D	$\mathbf{P}(H \mid D)$
h	d	0.95
$\neg h$	d	0.05
h	$\neg d$	0.25
$\neg h$	$\neg d$	0.75

Note that from the CPT's above, we can obtain CPT's for $\mathbf{P}(D \mid F)$ and $\mathbf{P}(F \mid D)$.

D	F	$\mathbf{P}(D \mid F)$
d	f	0.666
$\neg d$	f	0.334
d	$\neg f$	0.370
$\neg d$	$\neg f$	0.630

F	D	$\mathbf{P}(F \mid D)$
f	d	0.586
$\neg f$	d	0.414
f	$\neg d$	0.300
$\neg f$	$\neg d$	0.700

Note: Those are already given in the problem formulation, but try to derive them yourself as an exercise!

Solution:

(b) (ii) (Continued) Similarly, like in the exercise (a), we obtain by simply using the definitions provided in the class:

$$
\begin{aligned}
* \; EU(learn \mid f) &= \sum_d \mathrm{P}(d \mid f) U(d, learn) \\
&= \mathrm{P}(d \mid f) U(d, learn) + \mathrm{P}(\neg d \mid f) U(\neg d, learn) \\
&= 0.666 \cdot 1400 + 0.334 \cdot 800 = 1199.6 \\
* \; EU(sleep \mid f) &= \mathrm{P}(d \mid f) U(d, sleep) + \mathrm{P}(\neg d \mid f) U(\neg d, sleep) \\
&= 0.666 \cdot 0 + 0.334 \cdot 2000 = 668 \\
* \; EU(learn \mid \neg f) &= \mathrm{P}(d \mid \neg f) U(d, learn) + \mathrm{P}(\neg d \mid \neg f) U(\neg d, learn) \\
&= 0.370 \cdot 1400 + 0.630 \cdot 800 = 1022 \\
* \; EU(sleep \mid \neg f) &= \mathrm{P}(d \mid \neg f) U(d, sleep) + \mathrm{P}(\neg d \mid \neg f) U(\neg d, sleep) \\
&= 0.370 \cdot 0 + 0.630 \cdot 2000 = 1260
\end{aligned}
$$

(iii)

$$
\begin{aligned}
* \; MEU(\{f\}) &= \max_a EU(a \mid f) = \max(EU(sleep \mid f), EU(learn \mid f)) \\
&= \max(668, 1199.6) = 1199.6 \\
* \; MEU(\{\neg f\}) &= \max_a EU(a \mid \neg f) = \max(EU(sleep \mid \neg f), EU(learn \mid \neg f)) \\
&= \max(1260, 1022) = 1260
\end{aligned}
$$

We now obtain:

* Optimal Action($\{f\}$) = *learn*
* Optimal Action($\{\neg f\}$) = *sleep*