DATA 609: Homework 8

Modeling With Decision Theory

Aaron Grzasko

October 14, 2017

```r
# load libraries
library(dplyr)
library(knitr)
library(kableExtra)
library(Rgraphviz)
```

Page 347, Question 4

We have engaged in a business venture. Assume the probability of success is $P(s) = \frac{2}{5}$; further assume that if we are successful we make \$55,000, and if we are unsuccessful we lose \$1,750. Find the expected value of the business venture.

Given:

$$
\begin{aligned}
S_1 = \text{successful outcome} = & \quad \$55,000 \\
S_2 = \text{unsuccesful outcome} = & \quad -\$1,750 \\
P(S_1) = \frac{2}{5} & \\
P(S_2) = \frac{3}{5} &
\end{aligned}
$$

Solve:

$$
\begin{aligned}
E(S) &= \sum_{i=1}^{2} S_i P(S_i) \\
&= \$55,000 \times \frac{2}{5} - \$1,750 \times \frac{3}{5} \\
&= \$20,950
\end{aligned}
$$

Page 347, Question 6

Consider a firm handling concessions for a sporting event. The firm's manager needs to know whether to stock up with coffee or cola and is formulating policies for specific weather predictions. A local agreement restricts the firm to selling only one type of beverage. The firm estimates a \$1,500 profit selling cola if the weather is cold and a \$5,000 profit selling cola if the weather is warm. The firm also estimates a \$4,000 profit selling coffee if it is cold and a \$1,000 profit selling coffee if the weather is warm. The weather forecast says that there is a 30% of a cold front; otherwise, the weather will be warm. Build a decision tree to assist with the decision. What should the firm handling concessions do?

$$
\begin{aligned}
EV(\text{cola profit}) &= (\text{cola profit} \mid \text{cold}) \times P(\text{cold}) + (\text{cola profit} \mid \text{warm}) \times P(\text{warm}) \\
&= \$1,500 \times 0.3 + \$5,000 \times 0.7 \\
&= \$3,950
\end{aligned}
$$

$$
\begin{aligned}
EV(\text{coffee profit}) &= (\text{coffee profit} \mid \text{cold}) \times P(\text{cold}) + (\text{coffee profit} \mid \text{warm}) \times P(\text{warm}) \\
&= \$4,000 \times 0.3 + \$1,000 \times 0.7 \\
&= \$1,900
\end{aligned}
$$

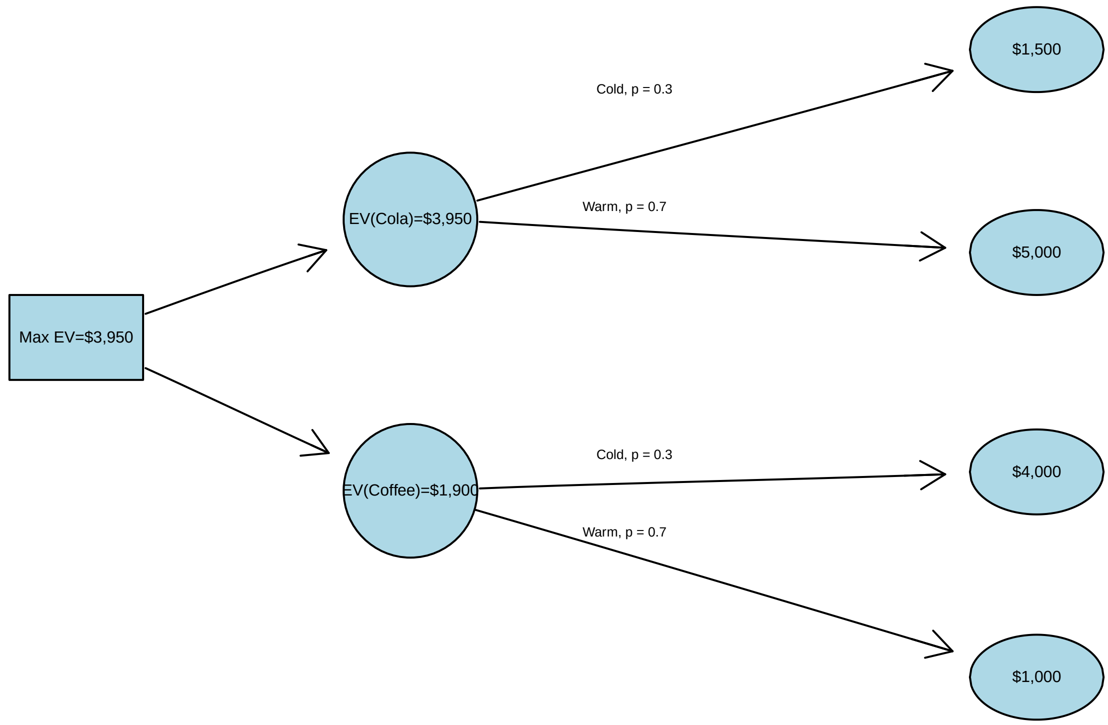
$$
Max\{EV(\text{cola profit}), EV(\text{coffee profit})\} = \$3,950
$$

Assuming the objective is to maximize the expected value of profit, the firm should stock up on cola, as this option results in an additional $2,050 in profit, on average, compared to coffee.

Below, we display the problem in the form of a decision tree:

```r
# nodes
node1<-"Max EV=$3,950"
node2<-"EV(Cola)=$3,950"
node3<-"EV(Coffee)=$1,900"
node4<-"$1,500"
node5<-"$5,000"
node6<-"$4,000"
node7<-"$1,000"
nodeNames<-c(node1,node2,node3,node4, node5,node6, node7)

# graph object
rEG <- new("graphNEL", nodes=nodeNames, edgemode="directed")

# edges
rEG <- addEdge(nodeNames[1], nodeNames[2], rEG, 1)
rEG <- addEdge(nodeNames[1], nodeNames[3], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[4], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[5], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[6], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[7], rEG, 1)

eAttrs <- list()
q<-edgeNames(rEG)

eAttrs$label <- c("", "","Cold, p = 0.3", "Warm, p = 0.7", "Cold, p = 0.3",
                  "Warm, p = 0.7")
names(eAttrs$label) <- c(q[1],q[2], q[3], q[4], q[5], q[6])
edgeAttrs<-eAttrs

# default attributes
attributes<-list(node=list(label="", fillcolor="lightblue", fontsize="15", shape="ellipse"),
 edge=list(color="black"),graph=list(rankdir="LR"))

# individual node attributes
nAttrs <- list()
nAttrs$label <- nodes(rEG)
```

```r
names(nAttrs$label) <- nodes(rEG)
nAttrs$shape <- c("Max EV=$3,950" = "box","EV(Cola)=$3,950"="circle","EV(Coffee)=$1,900"="circle")

# plot
plot(rEG, nodeAttrs=nAttrs, edgeAttrs=eAttrs, attrs=attributes)
```



Page 355, Question 3

The financial success of a ski resort in Squaw Valley is dependent on the amount of early snowfall in the fall and winter months. If the snowfall is greater than 40 inches, the resort always has a successful ski season. If the snow is between 30 and 40 inches,the resort has a moderate season, and if the snowfall is less than 30 inches, the season is poor, and the resort will lose money. The seasonal snow probabilities from the weather service are displayed in the following table with the expected revenue for the previous 10 seasons. A hotel chain has offered to lease the resort during the winter for $100,000. You must decide whether to operate yourself or lease the resort. Build a decision tree to assist in the decision.

	Financial return if we operate	Lease	Probability
Snow $\geq$ 40 in.	280,000	100,000	0.4
Snow 30 to 40 in.	100,000	100,000	0.2
Snow $\leq$ 30 in.	-40,000	100,000	0.4

First, let's solve by hand:

$$
\begin{aligned}
EV(\text{op profit}) &= (\text{op profit} \mid \text{heavy}) \times P(\text{heavy}) + (\text{op profit} \mid \text{mod}) \times P(\text{mod}) + (\text{op profit} \mid \text{light}) \times P(\text{light}) \\
&= \$280,000 \times 0.4 + \$100,000 \times 0.2 - \$40,000 \times 0.4 \\
&= \$116,000
\end{aligned}
$$

$$
\begin{aligned}
EV(\text{lease profit}) &= (\text{lease profit} \mid \text{heavy}) \times P(\text{heavy}) + (\text{lease profit} \mid \text{mod}) \times P(\text{mod}) + (\text{lease profit} \mid \text{light}) \times P(\text{light}) \\
&= \$100,000 \times (0.4 + 0.2 + 0.4) \\
&= \$100,000
\end{aligned}
$$

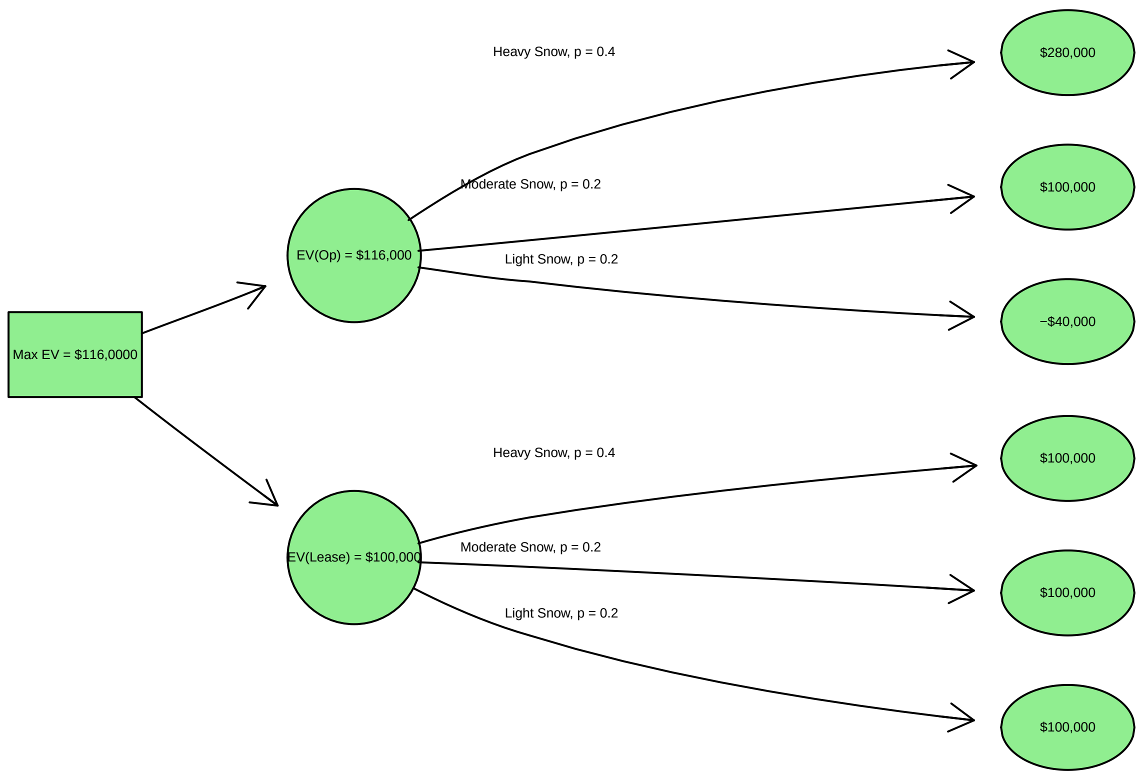
$$
Max\{EV(\text{op profit}), EV(\text{lease profit})\} = \$116,000
$$

Assuming our goal is to maximize expected value, we choose to operate the ski resort, as this option provides an additional $16,000 of profit, on average.

Now we'll construct the decision tree.

```r
# nodes
node1 <- "master"
node2 <- "opt1"
node3 <- "opt2"
node4 <- "out1.1"
node5 <- "out1.2"
node6 <- "out1.3"
node7 <- "out2.1"
node8 <- "out2.2"
node9 <- "out2.3"
nodeNames <- c(node1,node2,node3,node4,node5,node6,node7,node8,node9)

# graph object
rEG <- new("graphNEL", nodes=nodeNames, edgemode="directed")

# edges
rEG <- addEdge(nodeNames[1], nodeNames[2], rEG, 1)
rEG <- addEdge(nodeNames[1], nodeNames[3], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[4], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[5], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[6], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[7], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[8], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[9], rEG, 1)

eAttrs <- list()
q<-edgeNames(rEG)
eAttrs$label <- c("", "","Heavy Snow, p = 0.4", "Moderate Snow, p = 0.2", "Light Snow, p = 0.2",
                  "Heavy Snow, p = 0.4", "Moderate Snow, p = 0.2", "Light Snow, p = 0.2")
names(eAttrs$label) <- c(q[1],q[2], q[3], q[4], q[5], q[6], q[7], q[8])
edgeAttrs<-eAttrs

# default attributes
attributes<-list(node=list(label="", fillcolor="lightgreen", fontsize="15", shape="ellipse"),
 edge=list(color="black", fontsize = "15"),graph=list(rankdir="LR"))
```

```r
# individual node attributes
nAttrs <- list()
nAttrs$label <- c("Max EV = $116,0000","EV(Op) = $116,000","EV(Lease) = $100,000",
                  "$280,000","$100,000","-$40,000","$100,000","$100,000","$100,000")
names(nAttrs$label) <- nodes(rEG)
nAttrs$shape <- c(master = "box",opt1 ="circle",opt2 = "circle")

# plot
plot(rEG, nodeAttrs=nAttrs, edgeAttrs=eAttrs, attrs=attributes)
```



Page, 364, Question 3

A big private oil company must decide whether to drill in the Gulf of Mexico. It costs $1 Million to drill, and if oil is found its value is estimate at $6 million. At present, the oil company believes that there is a 45% change that oil is present. Before drilling begins, the big private oil company can hire a geologist for $100,000 to obtain samples and test for oil. There is only about a 60% chance that the geologist will issue a favorable report. Given that the geologist does issue a favorable report, there is an 85% chance that there is oil. Given an unfavorable report, there is a 22% chance that there is oil. Determine what the big private oil company should do.

Let

$$
\begin{aligned}
I &= \text{income} \\
G &= \text{geologist} \\
G' &= \text{no geologist} \\
F &= \text{favorable review} \\
F' &= \text{unfavorable review} \\
D &= \text{drill} \\
D' &= \text{no drill} \\
O &= \text{oil} \\
O' &= \text{no oil}
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G,F,D) &= (I|G,F,D,O) \times P(O|G,F,D) + (I|G,F,D,O') \times P(O'|G,F,D) \\
&= (\$6M - \$1M - \$0.1M) \times 0.85 + (\$0M - \$1M - \$0.1M) \times 0.15 \\
&= \$4.9M \times 0.85 - \$1.1M \times 0.15 \\
&= \$4M
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G,F,D') &= (I|G,F,D',O) \times P(O|G,F,D') + (I|G,F,D',O') \times P(O'|G,F,D') \\
&= (\$0M - \$0.1M) \times 0.85 + (\$0M - \$0.1M) \times 0.15 \\
&= -\$0.1M
\end{aligned}
$$

$$
\begin{aligned}
Max[EV(I,G,F)] &= Max\{EV(I|G,F,D), EV(I|G,F,D')\} \\
&= Max\{\$4M, -\$0.1M\} \\
&= \$4M
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G,F',D) &= (I|G,F',D,O) \times P(O|G,F',D) + (I|G,F',D,O') \times P(O'|G,F',D) \\
&= (\$6M - \$1M - \$0.1M) \times 0.22 + (\$0M - \$1M - \$0.1M) \times 0.78 \\
&= \$4.9M \times 0.22 - \$1.1M \times 0.78 \\
&= \$0.22M
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G,F',D') &= (I|G,F',D',O) \times P(O|G,F',D') + (I|G,F',D',O') \times P(O'|G,F',D') \\
&= (\$0M - \$0.1M) \times 0.22 + (\$0M - \$0.1M) \times 0.78 \\
&= -\$0.1M
\end{aligned}
$$

$$
\begin{aligned}
Max[EV(I|G,F')] &= Max\{EV(I|G,F',D), EV(I|G,F',D')\} \\
&= Max\{\$0.22M, \$0.1M\} \\
&= \$0.22M
\end{aligned}
$$

$$
\begin{aligned}
Max[EV(I|G)] &= Max[EV(I|G,F)] \times P(F|G) + Max[EV(I|G,F')] \times P(F'|G) \\
&= \$4M \times 0.6 + \$0.22M \times 0.4 \\
&= \$2.488M
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G',D) &= (I|G,D,O) \times P(O|G,D) + (I|G,D,O') \times P(O'|G,D) \\
&= (\$6M - \$1M) \times 0.45 - \$1M \times 0.55 \\
&= \$1.7M
\end{aligned}
$$

$$
\begin{aligned}
EV(I|G',D') &= (I|G,D',O) \times P(O|G,D') + (I|G,D',O') \times P(O'|G,D') \\
&= \$0M \times 0.45 + \$0M \times 0.55 \\
&= \$0M \\
Max[EV(I|G')] &= Max\{EV(I|G',D), EV(I|G',D')\} \\
&= Max\{\$1.7M, \$0M\} \\
&= \$1.7M
\end{aligned}
$$

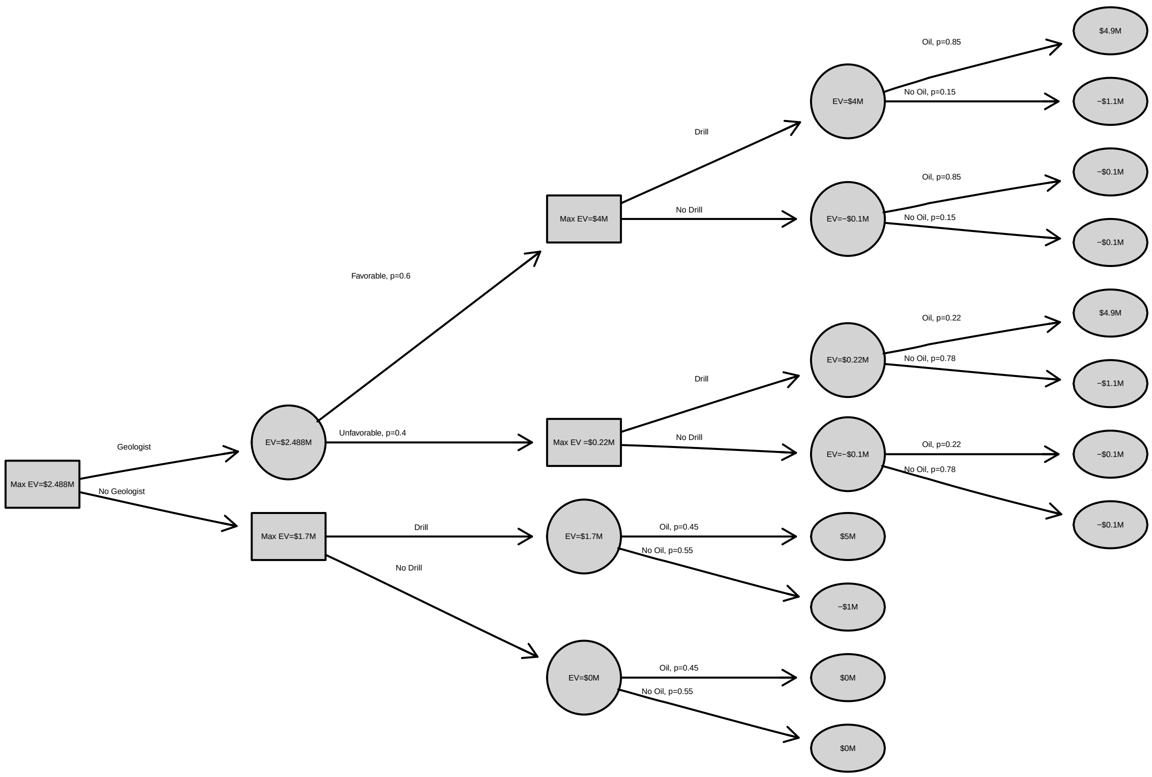
$$
\begin{aligned}
Max[EV(I)] &= Max\{Max[EV(I|G)], Max[EV[I|G']\} \\
&= Max\{\$2.488M, \$1.7M\} \\
&= \$2.488M
\end{aligned}
$$

To maximize expected value, the firm should hire the geologist. The firm should then drill for oil, regardless of the whether the outlook is favorable or unfavorable.

Below we recreate the problem as a decision tree.

```
# nodes
node1 <- "master"
node2 <- "geo"
node3 <- "noGeo"
node4 <- "geoFav"
node5 <- "geoNoFav"
node6 <- "geoFavDrill"
node7 <- "geoFavNoDrill"
node8 <- "geoNoFavDrill"
node9 <- "geoNoFavNoDrill"
node10 <- "noGeoDrill"
node11 <- "noGeoNoDrill"
node12 <- "geoFavDrillOil"
node13 <- "geoFavDrillNoOil"
node14 <- "geoFavNoDrillOil"
node15 <- "geoFavNoDrillNoOil"
node16 <- "geoNoFavDrillOil"
node17 <- "geoNoFavDrillNoOil"
node18 <- "geoNoFavNoDrillOil"
node19 <- "geoNoFavNoDrillNoOil"
node20 <- "noGeoDrillOil"
node21 <- "noGeoDrillNoOil"
node22 <- "noGeoNoDrillOil"
node23 <- "noGeoNoDrillNoOil"


nodeNames <- c(node1,node2,node3,node4,node5,node6,node7,node8,node9, node10,
               node11, node12, node13, node14, node15, node16, node17, node18,
               node19, node20, node21, node22, node23)

# graph object
rEG <- new("graphNEL", nodes=nodeNames, edgemode="directed")

# edges
rEG <- addEdge(nodeNames[1], nodeNames[2], rEG, 1)
rEG <- addEdge(nodeNames[1], nodeNames[3], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[4], rEG, 1)
rEG <- addEdge(nodeNames[2], nodeNames[5], rEG, 1)
rEG <- addEdge(nodeNames[4], nodeNames[6], rEG, 1)
rEG <- addEdge(nodeNames[4], nodeNames[7], rEG, 1)
rEG <- addEdge(nodeNames[5], nodeNames[8], rEG, 1)
rEG <- addEdge(nodeNames[5], nodeNames[9], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[10], rEG, 1)
rEG <- addEdge(nodeNames[3], nodeNames[11], rEG, 1)
rEG <- addEdge(nodeNames[6], nodeNames[12], rEG, 1)
rEG <- addEdge(nodeNames[6], nodeNames[13], rEG, 1)
rEG <- addEdge(nodeNames[7], nodeNames[14], rEG, 1)
rEG <- addEdge(nodeNames[7], nodeNames[15], rEG, 1)
rEG <- addEdge(nodeNames[8], nodeNames[16], rEG, 1)
rEG <- addEdge(nodeNames[8], nodeNames[17], rEG, 1)
```

```r
rEG <- addEdge(nodeNames[9], nodeNames[18], rEG, 1)
rEG <- addEdge(nodeNames[9], nodeNames[19], rEG, 1)
rEG <- addEdge(nodeNames[10], nodeNames[20], rEG, 1)
rEG <- addEdge(nodeNames[10], nodeNames[21], rEG, 1)
rEG <- addEdge(nodeNames[11], nodeNames[22], rEG, 1)
rEG <- addEdge(nodeNames[11], nodeNames[23], rEG, 1)



eAttrs <- list()
q<-edgeNames(rEG)
eAttrs$label <- c("Geologist", "No Geologist", "Favorable, p=0.6",
                  "Unfavorable, p=0.4",rep(c("Drill", "No Drill"),3),
                  rep(c("Oil, p=0.85", "No Oil, p=0.15"),2),
                  rep(c("Oil, p=0.22", "No Oil, p=0.78"),2),
                  rep(c("Oil, p=0.45", "No Oil, p=0.55"),2))
names(eAttrs$label) <- c(q[1],q[2], q[3], q[4], q[5], q[6], q[7], q[8], q[9],
                         q[10],q[11],q[12],q[13],q[14], q[15],q[16],q[17],
                         q[18],q[19],q[20],q[21],q[22])
edgeAttrs<-eAttrs

# default attributes
attributes<-list(node=list(label="", fillcolor="lightgray", fontsize="15", shape="ellipse"),
 edge=list(color="black", fontsize = "12"),graph=list(rankdir="LR"))

# individual node attributes
nAttrs <- list()
nAttrs$label <- c("Max EV=$2.488M","EV=$2.488M","Max EV=$1.7M","Max EV=$4M",
                  "Max EV =$0.22M","EV=$4M",
                  "EV=-$0.1M","EV=$0.22M","EV=-$0.1M","EV=$1.7M","EV=$0M",
                  "$4.9M","-$1.1M","-$0.1M", "-$0.1M","$4.9M","-$1.1M",
                  "-$0.1M", "-$0.1M","$5M", "-$1M", "$0M", "$0M")
names(nAttrs$label) <- nodes(rEG)
nAttrs$shape <- c(master = "box", geo = "circle", noGeo = "box", geoFav = "box",
                  geoNoFav = "box", geoFavDrill = "circle", geoFavNoDrill = "circle",
                  geoNoFavDrill = "circle", geoNoFavNoDrill = "circle",
                  noGeoDrill = "circle",noGeoNoDrill = "circle")

# plot
plot(rEG, nodeAttrs=nAttrs, edgeAttrs=eAttrs, attrs=attributes)
```



Page 373, Question 1

Give the following payoff matrix, show all work to answer parts a and b.

```r
# define states and outcomes under 3 alternatives
prob <- c(0.35,0.3,0.25,0.1)
a <- c(1100,900,400,300)
b <- c(850,1500,1000,500)
c <- c(700,1200,500,900)

# save to df
mydf <- data.frame(states = 1:4, prob=prob, A=a, B=b, C=c)

# display
cnames <- c("State", "Probability", "Alt A", "Alt B", "Alt C")
kable(mydf, col.names = cnames, format.args=list(big.mark = ",", scientific = F))
```

State	Probability	Alt A	Alt B	Alt C
1	0.35	1,100	850	700
2	0.30	900	1,500	1,200
3	0.25	400	1,000	500
4	0.10	300	500	900

(A): Which alternative do we choose if our criterion is to maximize the expected value?

```r
# expected value A
EV.a <- with(mydf, sum(A*prob))
EV.a
```

```
[1] 785
```

```r
# expected value B
EV.b <- with(mydf, sum(B*prob))
EV.b
```

```
[1] 1047.5
```

```r
# expected value C
EV.c <- with(mydf, sum(C*prob))
EV.c
```

```
[1] 820
```

```r
# max expected value
max(EV.a, EV.b, EV.c)
```

```
[1] 1047.5
```

Alternative B maximizes the expected value.

(B): Find the opportunity loss (regret) table and compute the expected opportunity loss (regret) for each alternative. What decision do you make if your criterion is to minimize regret?

Here is the regret matrix:

```r
regret.a <- c(0,600,600,600)
regret.b <- c(250,0,0,400)
regret.c <- c(400,300,500,0)
prob <- c(0.35,0.3,0.25,0.1)

mydf <- data.frame(rbind(regret.a, regret.b, regret.c))

names(mydf) <- c(1,2,3,4)
row.names(mydf) <- c('A','B','C')

kable(mydf,format.args=list(big.mark = ",", scientific = F))
```

	1	2	3	4
A	0	600	600	600
B	250	0	0	400
C	400	300	500	0

Here is the expected regret for each alternative:

```r
# expected regret A
EVReg.A <- sum(regret.a * prob)
EVReg.A
```

```
[1] 390
```

```r
# expected regret B
EVReg.B <- sum(regret.b * prob)
EVReg.B
```

```
[1] 127.5
```

```r
# expected reget C
EVReg.C <- sum(regret.c * prob)
EVReg.C
```

```
[1] 355
```

```r
# min expected regret
min(EVReg.A, EVReg.B, EVReg.C)
```

```
[1] 127.5
```

If the goal is to minimize expected regret, then alternative B should be selected.

References

- Official Rgraphviz documentation: https://www.bioconductor.org/packages/devel/bioc/vignettes/Rgraphviz/inst/doc/Rgraphviz.pdf
- Practical examples with Rgraphviz: http://www.harrysurden.com/wordpress/archives/292
- Installing Rgraphviz: https://stackoverflow.com/questions/18023300/is-rgraphviz-no-longer-available-for-r