MGTS7526 Assignment 2 - Risk Modelling Assignment Sheet

The total length of your assignment should not exceed eight (8) pages.

1. Horse Race (10 marks)

Let's assume that there is a race between two horses: Fleetfoot and Dogmeat, and you want to determine which horse to bet on. Fleetfoot and Dogmeat have raced against each other on twelve previous occasions, all two-horse races. Of these last twelve races, Dogmeat won five and Fleetfoot won the other seven. Therefore, all other things being equal, the probability of Dogmeat winning the next race can be estimated as 5/12 or 0.417 or 41.7%. However, on three of Dogmeat's previous five wins, it had rained before the race. It had rained only once on any of the days that he lost. On the day of the race in question, it is raining.

Construct a Bayesian network to show the probability of Dogmeat winning the race. Explain your Bayesian network and how you obtained your answer.

2. Meat Test (10 marks)

Minced meat purchased in the supermarket may be infected with bacteria. On average, infection occurs once in 600 packages of meat. A test with a *positive* or *negative* result can be used to test for infection. If the meat is *clean*, the test result will be *negative* in 499 out of 500 cases, and if the meat is *infected*, the test result will be *positive* in 499 out of 500 cases.

Construct a Bayesian Network to show the probability of a package of meat being *infected* given a *positive* test result. Explain your Bayesian network and how you obtained your answer.

3. Flower Breeding (20 marks)

You are a flower breeder. The plant you are breeding can either have red flowers or white flowers. You know that the colour of a flower depends on the genotype of the plant. The gene for red flowers (represented by R) is a dominant gene and the gene for white flowers (represented by r) in a recessive gene. Therefore, a plant with the genotype RR or Rr has red flowers, while a plant with the genotype rr has white flowers. Hence, the colour of a plant’s flowers is influenced by its genotype (as shown in Figure 1) and the probability of a plant having red or white flowers, given its genotype, is shown in Table 1.

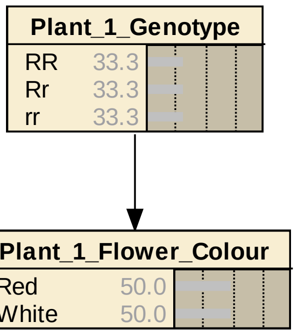

Figure 1: Diagram showing that a plant’s genotype influences the colour of its flowers.

Table1: Probability of a plant having red or white flowers given its genotype.

Genotype	Probability of flower colour (%)	
	Red	White
RR	100	0
Rr	100	0
rr	0	100

When breeding flowers you know that the genotype (and therefore flower colour) of an offspring is influenced by the genotype of its parents (as shown in Figure 2).

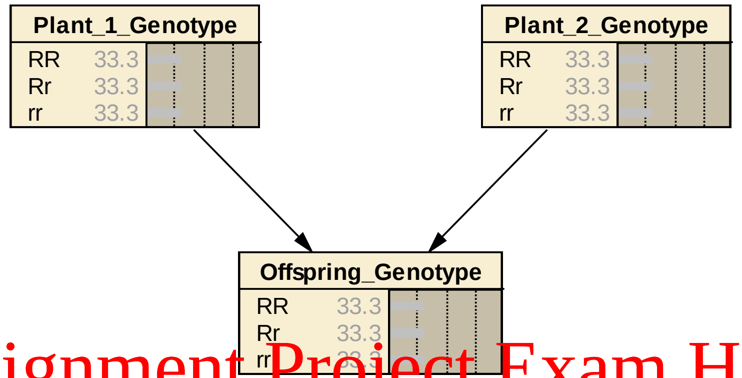

Figure 2: Diagram showing that the genotype of an offspring is influenced by the genotype of its parents.

You also know that the following parent crosses are possible:

- If two plants of genotype RR are mated, then the offspring will always be RR.
- If two plants of genotype rr are mated, then the offspring will always be rr.
- If a plant of genotype RR is mated with a plant of genotype Rr, then the offspring will always get an R from one parent and may get an R or an r from the other parent, which means it could be of genotype RR or Rr.
- If a plant of genotype RR is mated with a plant of genotype rr, then the offspring will always get an R from one parent and will always get an r from the other parent, which means it will always be of genotype Rr.
- If a plant of genotype Rr is mated with a plant of genotype Rr, then the offspring may get an R or r from one parent and an R or r from the other parent, which means it could be of genotype RR, Rr or rr.
- If a plant of genotype Rr is mated with a plant of genotype rr, then the offspring may get an R or r from one parent and will always get an r from the other parent, which means it could be of genotype Rr or rr.

For the above crosses, the probability of offspring being a particular genotype is given in Table 2.

Table 2: Probability of offspring genotypes given the genotypes of the parents

Parent Genotypes		Probability of Offspring Genotype (%)		
Parent 1	Parent 2	RR	Rr	rr
RR	RR	100	0	0
rr	rr	0	0	100
RR	Rr	50	50	0
RR	rr	0	100	0
Rr	Rr	25	50	25
Rr	rr	0	50	50

Finally, for plants with unknown parent genotypes, you know that the probability of them being genotype Rr is 50%, while the probability of them being of genotype RR or rr is 25%.

Now suppose you have two plants. The genotypes of their parents are unknown; however the flowers of both plants are red. You mate these two plants to produce a first generation offspring. This offspring is then mated with a third plant, with white flowers, to produce a second generation offspring.

Construct a Bayesian Network and use it to determine the probability that the second generation offspring will have red flowers? Explain your Bayesian network and how you obtained your answer.

4. Horse Stud (20 marks)

You are the manager of a horse stud. A colt called John has been found to suffer from a life-threatening hereditary disease caused by a recessive gene. The disease is so serious that John's parents, Henry and Irene, are taken out of the stud-breeding program. However, you still need to decide which of the remaining horses in the stud are likely to carry the disease-causing gene and therefore should be removed from the breeding program. You look through the stud records to retrace John's family tree (Table 1).

Table 1: John's family tree.

Mare	Stallion	Foal
Irene	Henry	John
Dorothy	Fred	Henry
Gwenn	Eric	Irene
Jill	Jack	Fred
Jill	Brian	Dorothy
Cecily	Brian	Eric
Cecily	Mike	Gwenn

You know that in order to have the disease, a horse must carry a double-recessive gene (aa). John is the only horse in the stud that has the disease so he can be the only horse of the genotype aa. Therefore, the remaining horses in the breeding program can either be carriers of the disease causing gene (aA) or pure (AA). You do some further research and find the probabilities of a foal being diseased (aa), a carrier (aA) or pure (AA), given the genotype of the father and mother (Table 2).

Table 2: Probability of a foal being diseased given the genotype of the father and mother.

		Foal		
Father	Mother	aa	aA	AA
aA	aA	25%	50%	25%
aA	AA	0%	50%	50%
AA	aA	0%	50%	50%
AA	AA	0%	0%	100%

We know that John is the only horse with the genotype aa, so for the other horses in the stud we can remove the probability of them being diseased (aa) and normalise the remaining probabilities so that they add to 100%. Hence, for the other horses in the stud, the probability of them being a carrier (aA) or pure (AA) given the genotype of the father and mother is shown in Table 3.

Table 3: Probability of a foal being a carrier or pure given the genotype of the father and mother.

		Foal	
Father	Mother	aA	AA
aA	aA	67%	33%
aA	AA	50%	50%
AA	aA	50%	50%
AA	AA	0%	100%

For the horses without a recorded father or mother, we know that the frequency of occurrence of the recessive gene is 1 in every 100 horses.

Construct a Bayesian Network to help you determine which horses in the stud are most likely to be carriers of the disease-causing gene and should be culled from the breeding program. You have one further piece of information to assist your decision – **Fred has previously been tested for the disease-causing gene and he is not a carrier (he is pure)**. Which horse(s) will you cull from the breeding program? Explain your Bayesian network and how you obtained your answer.