CS360 Homework 7– Solution

Genetic Algorithms

1) What are the advantages and disadvantages of carrying over the fittest two individuals to the next generation?

Answer:
Carrying over the fittest individuals (also known as 'elitism') tilts the genetic algorithm towards local search, because the best individuals stay around to (mostly) create more and more children very near to themselves in the genotype space. This could make the genetic algorithm more efficient, but also can cause it to more easily get caught in local optima.

2) What are the advantages and disadvantages of running genetic algorithms with only mutations and no crossovers? How about only crossovers and no mutations?

Answer:
Genetic algorithms without mutations cannot explore outside the possible points in space formed by choosing any combination of parameter settings in its original population. As individuals are weeded out, this combination reduces, ultimately to a single individual. For instance, if the individuals are represented as bit strings and the first bit of every individual is 0 in the first generation or becomes 0 (if all the 1's are eliminated), then the algorithm cannot generate solutions whose first bits are 1.

Genetic algorithms without crossovers perform purely random searches whose children are very similar to their parents. Crossovers help rapidly transfer high-performing parameters throughout the population, making genetic algorithms more efficient in problems where the parameters are at least somewhat independent of one another.

So, in general, mutations keep the search (somewhat) global and crossovers make it more efficient.

3) How would you encode a state if you were using a genetic algorithm to solve the Traveling Salesman Problem but only wanted to use a straightforward crossover operation that switches prefixes of both parents' encodings (that is, we randomly pick a cutoff point in the encoding, use the encoding of the first parent up to that cutoff point, and use the encoding of the second parent after that cutoff point)?

Answer:
We can, for example, assign an integer to each city, encoding its priority. These

priorities correspond to a unique tour, where the salesman visits the cities in the order of their priorities (breaking ties lexicographically). For this encoding, crossovers and mutations always create tours and thus valid solutions. Compare this encoding to one where, for each vertex, we keep the next vertex to visit. The crossover operator would have to be defined very carefully in order to generate valid solutions.

Probabilistic Reasoning

4) Consider the following joint probability table

PASSEXAM	WILDPARTY	
t	t	0.1
t	f	0.2
f	t	0.3
f	f	0.4

Calculate:

P(PASSEXAM)
Answer: 0.1 + 0.2 = 0.3

P($\neg$ WILDPARTY)
Answer: 0.2 + 0.4 = 0.6

P(PASSEXAM $\vee$ WILDPARTY)
Answer: 0.1 + 0.2 + 0.3 = 0.6

P(PASSEXAM $\Rightarrow$ WILDPARTY)
Answer: 0.1 + 0.3 + 0.4 = 0.8

P(PASSEXAM | WILDPARTY)
Answer: 0.1 / (0.1 + 0.3) = 0.25

P(WILDPARTY | PASSEXAM)
Answer: 0.1 / (0.1 + 0.2) = 0.33

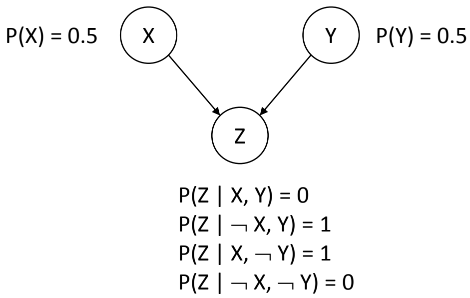
5) Remember that X and Y are called independent iff (= if and only if) $\mathrm{P}(X|Y) = \mathrm{P}(X)$ for all values of X and Y (that is, iff $\mathrm{P}(X = a|Y = b) = \mathrm{P}(X = a)$ for all values a that X can take on and all values b that Y can take on). Call X and Y conditionally independent given Z iff $\mathrm{P}(X|Y, Z) = \mathrm{P}(X|Z)$ for all values of X, Y and Z. (Note: $\mathrm{P}(X|Y, Z)$ means $\mathrm{P}(X|Y \wedge Z)$.)

Consider any X, Y and Z. Prove or disprove:

a) If X and Y are independent, are then X and Y necessarily conditionally independent given Z? (Assume that all conditional probabilities are well-defined.)

Answer:
We prove that the statement does not hold by giving a counter-example. Consider the following scenario where we flip two (unbiased) coins, a nickel and a

dime, and a bell rings if exactly one of them comes up heads (X: nickel coming up heads, Y: dime coming up heads, Z: bell ringing).



Then, X and Y are independent but X and Y are not guaranteed to be independent given Z. In fact, here they are not since, e.g.

$\mathrm{P}(X|Z) = \mathrm{P}(X, Z)/\mathrm{P}(Z) = 0.25/0.5 = 0.5$

$\mathrm{P}(X|Y, Z) = \mathrm{P}(X, Y, Z)/\mathrm{P}(Y, Z) = 0/0.25 = 0$

with

$\mathrm{P}(X, Y, Z) = \mathrm{P}(X)\ \mathrm{P}(Y)\ \mathrm{P}(Z|X, Y) = 0.5 \times 0.5 \times 0 = 0$

$\mathrm{P}(X, \neg Y, Z) = \mathrm{P}(X)\ \mathrm{P}(\neg Y)\ \mathrm{P}(Z|X, \neg Y) = 0.5 \times 0.5 \times 1 = 0.25$

$\mathrm{P}(\neg X, Y, Z) = \mathrm{P}(\neg X)\ \mathrm{P}(Y)\ \mathrm{P}(Z|\neg X, Y) = 0.5 \times 0.5 \times 1 = 0.25$

$\mathrm{P}(\neg X, \neg Y, Z) = \mathrm{P}(\neg X)\ \mathrm{P}(\neg Y)\ \mathrm{P}(Z|\neg X, \neg Y) = 0.5 \times 0.5 \times 0 = 0$

$\mathrm{P}(Y, Z) = \mathrm{P}(X, Y, Z) + \mathrm{P}(\neg X, Y, Z) = 0 + 0.25 = 0.25$

$\mathrm{P}(X, Z) = \mathrm{P}(X, Y, Z) + \mathrm{P}(X, \neg Y, Z) = 0 + 0.25 = 0.25$

$\mathrm{P}(Z) = \mathrm{P}(X, Y, Z) + \mathrm{P}(X, \neg Y, Z) + \mathrm{P}(\neg X, Y, Z) + \mathrm{P}(\neg X, \neg Y, Z)$

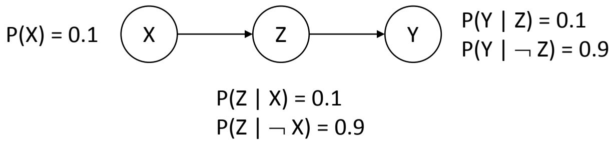
$= 0 + 0.25 + 0.25 + 0 = 0.5$

Of course, one does not need to use Bayesian networks. For example, one can use the joint probability table specified by the Bayesian network directly to prove that X and Y are independent but X and Y are not independent given Z.

X	Y	Z	
t	t	t	$0.5 \times 0.5 \times 0 = 0$
t	t	f	$0.5 \times 0.5 \times 1 = 0.25$
t	f	t	$0.5 \times 0.5 \times 1 = 0.25$
t	f	f	$0.5 \times 0.5 \times 0 = 0$
f	t	t	$0.5 \times 0.5 \times 1 = 0.25$
f	t	f	$0.5 \times 0.5 \times 0 = 0$
f	f	t	$0.5 \times 0.5 \times 0 = 0$
f	f	f	$0.5 \times 0.5 \times 1 = 0.25$

b) If X and Y are conditionally independent given Z, are then X and Y necessarily independent? (Assume that all conditional probabilities are well-defined.)

Answer:
We prove that the statement does not hold by giving a counter-example. Consider



Then, X and Y are independent given Z but X and Y are not guaranteed to be independent. In fact, here they are not since, e.g.

$\mathrm{P}(X) = 0.1$

$\mathrm{P}(X|Y) = \mathrm{P}(X, Y)/\mathrm{P}(Y) = 0.082/0.244 \sim 0.34$

with

$\mathrm{P}(X, Y) = \mathrm{P}(X, Y, Z) + \mathrm{P}(X, Y, \neg Z)$

$= \mathrm{P}(X)\ \mathrm{P}(Z|X)\ \mathrm{P}(Y|Z) + \mathrm{P}(X)\ \mathrm{P}(\neg Z|X)\ \mathrm{P}(Y|\neg Z)$

$= 0.1 \times 0.1 \times 0.1 + 0.1 \times 0.9 \times 0.9 = 0.082$

$\mathrm{P}(Y) = \mathrm{P}(X, Y) + \mathrm{P}(\neg X, Y)$

$= \mathrm{P}(X, Y) + \mathrm{P}(\neg X)\ \mathrm{P}(Z|\neg X)\ \mathrm{P}(Y|Z) + \mathrm{P}(\neg X)\ \mathrm{P}(\neg Z|\neg X)\ \mathrm{P}(Y|\neg Z)$

$= 0.082 + 0.9 \times 0.9 \times 0.1 + 0.9 \times 0.1 \times 0.9 = 0.244$

Of course, one does not need to use Bayesian networks. For example, one can use the joint probability table specified by the Bayesian network directly to prove that X and Y are independent given Z but X and Y are not independent.

X	Z	Y	
t	t	t	$0.1 \times 0.1 \times 0.1 = 0.001$
t	t	f	$0.1 \times 0.1 \times 0.9 = 0.009$
t	f	t	$0.1 \times 0.9 \times 0.9 = 0.081$
t	f	f	$0.1 \times 0.9 \times 0.1 = 0.009$
f	t	t	$0.9 \times 0.9 \times 0.1 = 0.081$
f	t	f	$0.9 \times 0.9 \times 0.9 = 0.729$
f	f	t	$0.9 \times 0.1 \times 0.9 = 0.081$
f	f	f	$0.9 \times 0.1 \times 0.1 = 0.009$

6) Prove that $\mathrm{P}(X|Y, Z) = \mathrm{P}(X|Z)$ for all values of X, Y and Z iff $\mathrm{P}(Y|X, Z) = \mathrm{P}(Y|Z)$ for all values of X, Y and Z. (Assume that all conditional probabilities are well-defined.)

Answer:

$\mathrm{P}(X|Y, Z) = \mathrm{P}(X|Z) \Leftrightarrow$

$\mathrm{P}(X, Y, Z)/\mathrm{P}(Y, Z) = \mathrm{P}(X, Z)\ /\ \mathrm{P}(Z) \Leftrightarrow$

$\mathrm{P}(X, Y, Z)/\mathrm{P}(X, Z) = \mathrm{P}(Y, Z)\ /\ \mathrm{P}(Z) \Leftrightarrow$

$\mathrm{P}(Y|X, Z) = \mathrm{P}(Y|Z)$

Prove that $\mathrm{P}(X|Y,Z) = \mathrm{P}(X|Z)$ for all values of X, Y and Z iff $\mathrm{P}(X,Y|Z) = \mathrm{P}(X|Z)\ \mathrm{P}(Y|Z)$ for all values of X, Y and Z. (Assume that all conditional probabilities are well-defined.)

$\mathrm{P}(X,Y|Z) = \mathrm{P}(X|Z)\ \mathrm{P}(Y|Z) \Leftrightarrow$

$\mathrm{P}(X,Y,Z)/\mathrm{P(Z)} = (\mathrm{P}(X,Z)/\mathrm{P}(Z))(\mathrm{P}(Y,Z)/\mathrm{P}(Z)) \Leftrightarrow$

$\mathrm{P}(X,Y,Z) = \mathrm{P}(X,Z)\mathrm{P}(Y,Z)/\mathrm{P}(Z) \Leftrightarrow$

$\mathrm{P}(X,Y,Z)/\mathrm{P}(Y,Z) = \mathrm{P}(X,Z)\ /\ \mathrm{P}(Z) \Leftrightarrow$

$\mathrm{P}(X|Y,Z) = \mathrm{P}(X|Z)$

7) Assume that TEST1 and TEST2 are conditionally independent given FLU. What is the probability of FLU (being true) given TEST1 and TEST2 (being true) if

P(FLU) = 0.1

P(TEST1 | FLU) = 0.9

P(TEST1 | ¬ FLU) = 0.2

P(TEST2 | FLU) = 0.8

P(TEST2 | ¬ FLU) = 0.1

Answer:
We apply Bayes' rules to solve the problem. Since TEST1 and TEST2 are conditionally independent given FLU, it holds that:

P(TEST1, TEST2 | FLU) = P(TEST1 | FLU) P(TEST2 | FLU) = 0.72

P(TEST1, TEST2 | ¬ FLU) = P(TEST1 | ¬ FLU) P(TEST2 | ¬ FLU) = 0.02

We then proceed as follows to find P(FLU | TEST1, TEST2):

P(TEST1, TEST2, FLU) = P(TEST1, TEST2 | FLU) P(FLU) = 0.72 × 0.1 = 0.072

P(TEST1, TEST2, ¬ FLU) = P(TEST1, TEST2 | ¬ FLU) P(¬ FLU) = 0.02 × (1-0.1) = 0.018

P(FLU | TEST1, TEST2) = P(TEST1, TEST2, FLU) / (P(TEST1, TEST2, FLU) + P(TEST1, TEST2, ¬ FLU)) = 0.072/(0.072 + 0.018) = 0.8

So, if both tests are true, there is an 80% chance of having the flu.

8) a) Is it possible that $\mathrm{P}(X|Y) + \mathrm{P}(X|\neg Y) \neq 1$? Why or why not?

Answer:
Yes. We can use the first question as a counter-example:

P(PASSEXAM | WILDPARTY) = 0.1 / (0.1 + 0.3) = 0.25

P(PASSEXAM | ¬ WILDPARTY) = 0.2 / (0.2 + 0.4) = 0.33

b) Is it possible that $\mathrm{P}(X|Y) + \mathrm{P}(\neg X|Y) \neq 1$? Why or why not?

Answer:
No. The reason is that, given Y, X either happens or does not happen. There is no third alternative, therefore, $\mathrm{P}(X|Y) + \mathrm{P}(\neg X|Y) = 1$. Formally:

$$\mathrm{P}(X|Y) + \mathrm{P}(\neg X|Y) = \mathrm{P}(X,Y)/\mathrm{P}(Y) + \mathrm{P}(X,\neg Y)/\mathrm{P}(Y) = \mathrm{P}(Y)/\mathrm{P}(Y) = 1$$

9) a) Sven rolled 1 fair die. What's the probability of him having rolled a 1 given that he rolled an odd number? What's the probability of him having rolled a 1 given that he rolled an even number?

Answer:
1/3 and 0, respectively. For example,

P(ONE | ODD) = P(ONE, ODD)/P(ODD) = P(ONE)/P(ODD) = (1/6)/(1/2) = 1/3

P(ONE | EVEN) = P(ONE, EVEN)/P(EVEN) = 0/(1/2) = 0

b) Sven rolled 2 fair dice, a red one and a green one. What's the probability of his total score being 4? What's the probability of him having rolled a 2 with the red die given that his total score was 4 or higher?

Answer:
Sven could have rolled a 1-3, 2-2 or 3-1 to get a total score of 4. There are 36 different possible combinations he can roll with two dice and all are equally likely, so the probability of rolling a total of 4 is $3/36 = 1/12$.

In the second case, since Sven's total score was 4 or higher, he could not have rolled a 1-1, 1-2 or 2-1. So, there are only 33 possible ways he could have rolled the dice with a total score of 4 or higher. Among these 33, only 5 of them can have the red die roll a 2 (2-2, 2-3, 2-4, 2-5 and 2-6). Therefore, the answer is 5/33.

10) Three random variables X, Y and Z are called mutually independent iff $\mathrm{P}(X,Y,Z) = \mathrm{P}(X)\ \mathrm{P}(Y)\ \mathrm{P}(Z)$ for all values of X, Y and Z. Assume that X and Y are independent, X and Z are independent and Y and Z are independent. Are X, Y and Z then necessarily mutually independent? Why or why not?

Answer:
No. Consider the following joint probability table as a counter-example:

X	Y	Z	
t	t	t	0
t	t	f	0.25
t	f	t	0.25
t	f	f	0
f	t	t	0.25
f	t	f	0
f	f	t	0
f	f	f	0.25

In this example, there are four possibilities: Either exactly two of X, Y and Z are true, or none of them are. Observe that, $\mathrm{P}(X) = 0.25 + 0.25 = 0.5$, $\mathrm{P}(\neg X) = 1 - \mathrm{P}(X) = 0.5$ and $\mathrm{P}(X|Y) = \mathrm{P}(\neg X|Y) = \mathrm{P}(X|\neg Y) = \mathrm{P}(\neg X|\neg Y) = 0.25/(0.25 + 0.25) = 0.5$. Therefore, X and Y are independent. We can also show that X, Y and Z are all pairwise independent, using similar (symmetric) arguments. However, they are not mutually independent because $\mathrm{P}(X)\ \mathrm{P}(Y)\ \mathrm{P}(Z) = 0.125 \neq \mathrm{P}(X, Y, Z) = 0$.

11) In the worst case, how many probabilities does one need to specify at most to specify the joint probability table of 10 Boolean random variables? How does the answer change if the 10 Boolean variables are mutually independent?

Answer:
There are 2^{10} different interpretations for 10 Boolean random variables. Since the probabilities have to sum to 1, we can omit any one of the entries in a given joint probability table, making the worst case require us to specify $2^{10} - 1$ entries.

If the variables are mutually independent, we only need to specify the probability of being true for each random variable, since any joint probability can be computed by multiplication. Therefore, we need only 10 entries.

12) (Courtesy of Russell and Norvig) After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease and that the test is 99% accurate (that is, the probability of testing positive given that you have the disease is 0.99 as is the probability of testing negative given that you don't have the disease). The good news is that this is a rare disease, striking only one in 10,000 people. Why is it good news that the disease is rare? What are the chances that you actually have the disease?

Answer:
Let D denote the probability of having the disease, and T denote the probability of testing positive for the disease. The problem gives us $\mathrm{P}(D) = 0.0001$, $\mathrm{P}(T|D) = 0.99$ and $\mathrm{P}(\neg T|\neg D) = 0.99$, and wants us to find $\mathrm{P}(D|T)$. This can be computed with Bayes rule as follows:

$\mathrm{P}(D, T) = \mathrm{P}(D)\ \mathrm{P}(T|D) = 0.0001 \times 0.99 = 0.000099 \sim 0.0001$

$\mathrm{P}(\neg D, T) = \mathrm{P}(\neg D)\ \mathrm{P}(T|\neg D) = (1 - \mathrm{P}(D))\ (1 - \mathrm{P}(\neg T|\neg D)) = 0.9999 \times 0.01 = 0.009999 \sim 0.01$

$\mathrm{P}(D|T) = \mathrm{P}(D, T)\ /\ (\mathrm{P}(D, T) + \mathrm{P}(\neg D, T)) = 0.0001\ /\ (0.0001 + 0.01) \sim 0.01$.

So, there is only $\sim$ 1% chance of having the disease, given that the test was positive.