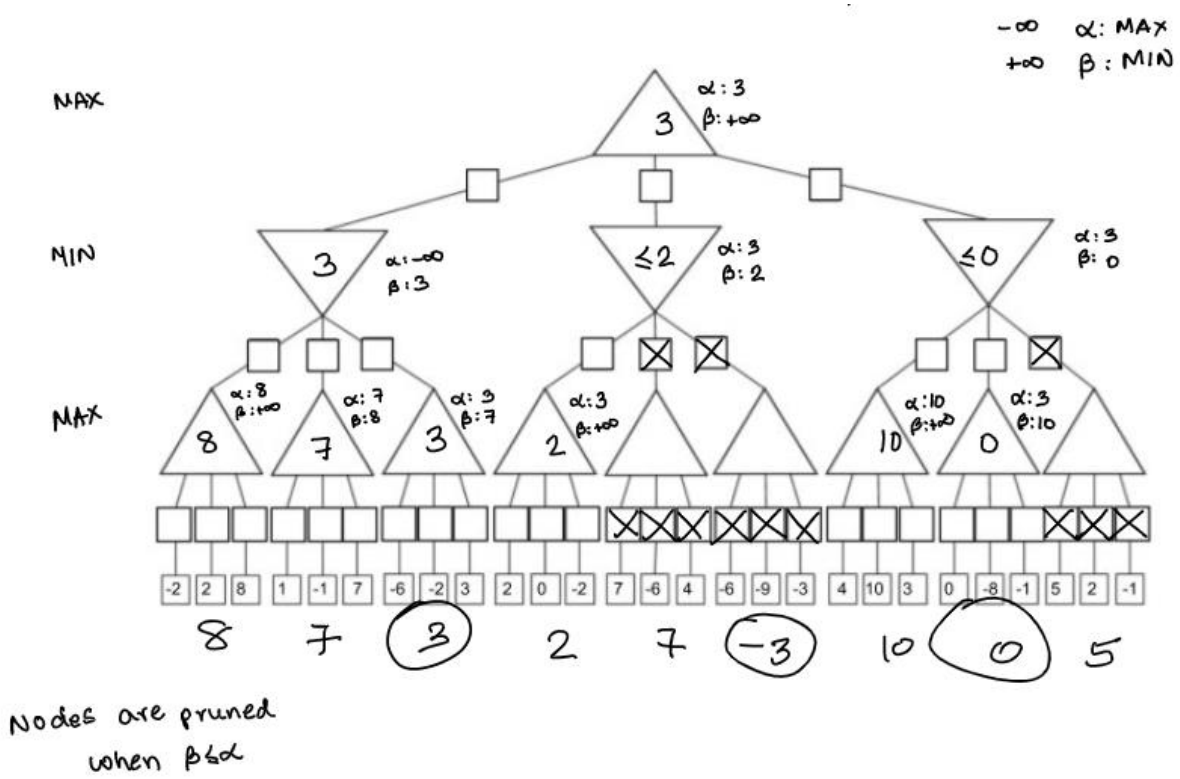
Midterm Spring 2020 Solutions

Question 1

A) There should be 12 nodes pruned, and the resulting tree should look something like this:



We allowed ranges or exact values for node scores due to the wording being confusing, but invalid scores were not allowed.

B) There were multiple correct tree configurations, but the best-case number of nodes pruned is 20.

C) There were multiple correct tree configurations, but the worst-case number of nodes pruned is 0.

Question 2

A) The optimal order of exchanges is **[ISBL, UGX, AUD, BTC, STNR]**.

B) The value of Maks's Starner Bucks is **9228.345375**. This is calculated by the following exchanges:

1. ISBL -> UGX: Cost is (10000) * (0.01) = 100. Remaining wealth is 9,900.
2. UGX -> AUD: Cost is (9,900) * (0.0023) = 22.77. Remaining wealth is 9877.23
3. AUD -> BTC: Cost is (9877.23) * (0.005) = 49.38615. Remaining wealth is 9827.84385.
4. BTC -> STNR: Cost is (9827.84385) * (0.0610) = 599.49847485. Remaining wealth is 9228.34537515.

Following the rounding rules of the exam, the final wealth in Isbellions is 9228.345375

C) 'Serbankia' is first added to the frontier with a cost of 124+244+135.464571013974 = **503.464571**

This is because 'Pandel' is the first node expanded in the search from Isbelland to Starneria, and Serbankia is first discovered as a neighbor of Pandel, and added to the frontier with the associated cost.

D) Using the weights in the PDF, 'Serbankia' is popped from the frontier with a cost of 153+211+135.464571013974 = **499.464571.** If you used the updated weights in the clarification thread, you'll end up with an alternative correct answer of **500.464571**

Even though Serbankia is initially added to the frontier with a cost of near 503, we re-discover Serbankia when we expand Thrunway. In doing so, we find that we have a better way of getting to the node, and update Serbankia's cost and associated path in the frontier accordingly. This cost will hold until Serbankia is popped from the frontier.

E) No, you cannot terminate the search. This doesn't hold in the A* algorithm, and one can see that the optimal path from Isbelland to Starneria does not include Essastan.

F) A heuristic is admissible if it never overestimates the cost between two nodes. Given that d(x) is an admissible heuristic, all the heuristics that maintain h(x) <= d(x) are admissible. There was some confusion around the logarithmic heuristic with the cost of a node to itself giving log(0), which is undefined. With this, we're ignoring this option within the grading rubric and selecting this answer will neither be penalized nor rewarded. With this in mind, the admissible heuristics are:

- **d(x) - 100**
- **1/2 * d(x)**

Partial Credit was awarded throughout the question if you made an error in rounding, got some of the multiple choice correct, and got some of the ideal currency exchanges correct.

Question 3

Part A

Please find below the solution to the Genetic Algorithms question in the midterm.

A.1

Specimen	Chromosome	Cost	Fitness Score
Network 1	{12, 12, 16, 20, 24}	4969.414502044605	0.000201
Network 2	{24, 24, 12, 12, 16}	4420.401985675268	0.000226
Network 3	{16, 20, 24, 12, 20}	4884.614328856155	0.000205
Network 4	{24, 12, 20, 16, 16}	4575.40615014309	0.000219

Partial Points:

- 0.5 for each entry. 0.25 deducted if rounding is wrong.
- 0.5 deducted overall if cost is reported.

A.2

Network 2

No partial points

A.3

Answers both in decimal and percentage format are accepted.

Specimen	Fitness Score	Probability to be a Parent	Probability in %
Network 5	0.000316	0.221599	22.159888
Network 6	0.000374	0.262272	26.227209
Network 7	0.000389	0.272791	27.279102
Network 8	0.000347	0.243338	24.333801

Partial Points: 0.5 for each entry. 0.25 deducted if rounding is wrong.

A.4

{16, 12, 20, 16, 12} and {20, 12, 20, 16, 24}

Partial Points: 1 point for each correct child. 0.25 point deducted for every incorrect answer

A.5

Accepting both D and B, D.

Partial Points:

- Only B: 0.5 points
- 0.5 point deducted for each incorrect answer

A.6

{16,16,24,16,20}

Partial points: 1 point if only the mutated gene is reported. No partial points otherwise.

I hope you liked the question. This question is motivated from the below paper with some parts slightly modified and simplified for the purpose of the exam.

Vairavamoorthy, Kalanithy, and Mohammed Ali. "Optimal design of water distribution systems using genetic algorithms." *Computer-Aided Civil and Infrastructure Engineering* 15, no. 5 (2000): 374-382.

Part B

Please find below the table, filled with the expected values.

	Task							Total Time Taken (hours)	Energy	ΔEnergy	Temp-erature	Probability of Acceptance
	T	W	S	G	A	B	End Time					
Start Time of Task	7:00	8:00	10:00	10:30	12:30	14:00	14:30	7.5	0.133333	-	0.07	1
	9:00	10:00	11:30	12:00	14:00	15:30	17:00	8	0.125	-0.008333	0.06	0.870330
	15:00	15:30	17:30	18:00	20:00	21:30	23:00	8	0.125	0	0.05	1
	17:00	17:30	19:30	20:30	00:00	01:00	01:30	8.5	0.117647	-0.007353	0.04	0.832081
	5:00	5:30	6:30	7:30	10:00	12:30	13:00	8	0.125	0.007353	0.03	1
	13:00	13:30	14:30	15:00	17:30	20:00	21:00	8	0.125	0	0.02	1
	11:00	11:30	13:00	13:30	16:00	18:30	19:30	8.5	0.117647	-0.007353	0.01	0.479362

Some random entries from the table have been graded (not all the entries).

For parts Q.3.B.1 and Q.3.B.2, all that is needed is values directly from the table.

(EDIT: Anyone who wrote the number incorrectly in their answer but has it correct in the table can submit a regrade request to get credit. This helps if you have transcription errors)

Please find the grading scheme below.

Q.3.B.1. Given that Alice starts at 9:00, answer the following questions:

Q.3.B.1a. What time does she start taking a shower? **(0.5 points)**

Ans: 11:30

(No partial credit)

Q.3.B.1b. What time does she finish all her tasks? **(0.5 points)**

Ans: 17:00

(No partial credit)

Q.3.B.1c. How many hours did she take to do all her tasks? **(0.5 points)**

Ans: 8

(No partial credit)

Q.3.B.1d. What is ΔEnergy for this start time? **(0.5 points)**

Ans: -0.008333

(-0.2 for incorrect sign)

Q.3.B.1e. What is the probability of acceptance for this start time? **(0.5 points)**

Ans: 0.870330

(-0.2 for incorrect sign)

Q.3.B.2. Given that Alice starts at 11:00, answer the following questions:

.

Q.3.B.2a. What time does she finish buying groceries from A-Star Supermarket? **(0.5 points)**

Ans: 18:30

(No partial credit)

Q.3.B.2b. What time does she finish all her tasks? **(0.5 points)**

Ans: 19:30

(No partial credit)

Q.3.B.2c. How many hours did she take to do all her tasks? **(0.5 points)**

Ans: 8.5

(No partial credit)

Q.3.B.2d. What is ΔEnergy for this start time? **(0.5 points)**

Ans: -0.007353

(-0.2 for incorrect sign)

Q.3.B.2e. What is the probability of acceptance for this start time? **(0.5 points)**

Ans: 0.479362

(-0.2 for incorrect sign)

Several answers did not follow the rounding rules of the question, but full credit has been given for minor rounding errors. Please be sure to read and follow all the instructions in the final exam.

Feel free to post below if there are any questions.

Question 4

Q.4.A.1

Expected Answer:

Jack (ex: requests going to island)

Jane (ex: requests going to G)

John (ex: cannot go to A)

Also accepting: [Jack, Jane, John, and Jamie] (with the assumption that students applied Note#2 here)

Partial credit: -⅓ for each missing person

Q.4.A.2

	A	B	C	D	E	F	G	H	I
Jack	O	X	O	X	O	X	O	O	O
Jane	O			O		O	**X**		O
John		O	O		O	O	X	X	O
Jamie	X	X	X	X	X	X	X	X	X

There are 36 cells in the table, and 1 was filled-in already. 3/35 points were subtracted for each incorrect cell in the submissions (3 points * 1/35).

Only the information in the client descriptions was needed for this question. The table itself indirectly represents the domains for each client.

An X was placed for all cases in which a client made a request that could be mapped to a region without other restrictions. If multiple regions could be used to satisfy a request, an X was to be placed on all that applied. Here are the explanations for all Xs:

Jack requests coastal regions (B, D, F)

Jack requests island (F)

Jane requests G

John requests shopping (A, D, G, H), but A, D are excluded due to his restrictions

Jamie requests northern regions (A, C, E, G, I)

Jamie requests southern regions (B, D, F, H)

A cell was to be left blank if the client had restrictions preventing him/her from traveling. Here are the explanations for the blanks:

Jane cannot go to hot regions (B, H)

Jane cannot go to C, E

John cannot go to A, D

If a client did not make a request to go to a region, but also had no restrictions from visiting, an O was to be placed. After filtering through the requests and restrictions:

Jack - A, C, E, G, H, I

Jane - A, D, F, I

John - B, C, E, F, I

(No partial credit was given if the table was left blank.)

Q.4.A.3.a

Answer: No

Explanation: The whole group must travel together and have all requests met. This is not possible for a few reasons:

- Jack cannot visit at least two coastal regions
- F is requested but is inaccessible (because B and D are not options)
- Jamie cannot visit at least two southern regions

(No partial credit)

Q.4.A.3.b

Expected answer: GI

Also accepting IG and IGI

Explanation: Based on the table from Q.4.A.2, regions F, G, I are options, but since F is not accessible, the group could not visit F. Note that this question was still graded even if "Yes" was answered in Q.4.A.3.a.

No credit was given if any region outside of {F, G, I} was included in the answer.

No credit was given if G and I were not present in the answer

Partial credit was given for the following cases:

- -0.25 if F was included
- -0.25 if any region (other than I) was present more than once

Q.4.A.4.a

Answer: Yes

Explanation: The whole group must travel together and have all requests met. Since B and H are now options based on the new information, this is now possible. The inclusion of B addresses the issues mentioned above for Q.4.A.3.

(No partial credit)

Q.4.A.4.b

Expected answer: IBFBGI

Also accepting: IGBFBI

Explanation: Now with the new information, regions B, F, G, H, I are options. Using these regions alphabetically to satisfy the constraints, all requests will be met without needing to visit H. Thus, including H would not be keeping the itinerary to a minimum. Note that this question was still graded even if "No" was answered in Q.4.A.4.a.

No credit was given if any region outside of {B, F, G, H, I} was included in the answer.

No credit was given if B, F, G, and I were not present in the answer

Partial credit was given for the following cases:

- -0.25 if H was included (itinerary not kept to minimum)
- -0.25 if any region (other than B or I) was present more than once (itinerary not kept to minimum)
- -0.25 if the itinerary does not contain the subsequence "BFB" (implying either that the itinerary has no plan to get to F, or the itinerary gets stuck once at F) (constraint from Note#1 not applied)
- -0.25 if the itinerary did not start and end with Region I (constraint from Note#2 not applied)

Part B

Q.4.B.1

Expected Answer:

No marinara sauce OR No alfredo sauce

Explanation: "No marinara sauce" and "No alfredo sauce" are both per slice constraints which make them unary constraints.

Notes: Some students marked both answers manually and these solutions were graded as correct. Since their was some confusion because of the question grammar conflicting with the radio buttons, we decided to give partial credit to students who answered with a combination of at least one of the correct answers.

Q.4.B.2

Expected Answer:

No adjacent slice can be meat

Explanation: Of all the possible answer choices, this constraint most directly relates a slice of pizza to its immediate neighbors. This constraint of different from each of its neighbors thus makes it a binary constraint.

Notes: Since their was some confusion because of the question grammar conflicting with the radio buttons, we decided to give partial credit to students who answered with a combination which contained the correct answer.

Q.4.B.3

Expected Answer:

Slice Number	Topping	Sauce
1	P	M
2	V	M
3	V	M
4	p	M
5	IS	M
6	IS	M
7	C	A
8	C	A

Explanation: Since the machine starts at slice 1, starts with vegetable toppings and can only switch between vegetable toppings and meat toppings once, we know the first slice needs to be either pineapple or vegetable. Since we now know that the first four slices will be either pineapple or vegetable, we know that the last four slices will be meat containing slices. Since the vegetable slices can't be adjacent to meat slices, and since slices 8 and 5 will be meat, slices 1 and 4 must be pineapple to shield the vegetable slices from the meat slices. This means slices 2 and 3 will be vegetable toppings. Next we deal with the constraint of the machine only being able to switch between sauces once. All the slices except the chicken alfredo slices can have marinara sauce, so to deal with the sauce switching constraint we just have the alfredo sauce slices last. This means slices 7 and 8 will be chicken topping with alfredo sauce. Since Jane refuses to eat alfredo sauce, this means that slices 5 and 6 need to have marinara sauce and their topping will be Italian Sausage.

Notes: For this question, the answer provided above is the singular correct answer. Since there are 16 fields for the question and the question is worth 4 points, we subtracted .25 points for every field that was different than the answer provided above.

Question 5

Section A:

5.A.1. If the test turns out to be positive, what probability should the doctor assign to the event that cancer is present?

Answer: 0.00197628

[No partial marking, but full credits are given if its mentioned in percentage or rounded]

5.A.2. Why is the probability calculated in the previous question so low (less than 0.5) despite the fact that 99 percent of cases in which cancer was present the test was positive?

Answer: C

[Partial credits are given]

5.A.3. Among the results where the test is positive, what is the false-positive rate (i.e., what is the percentage of positives that are due to error)?

Answer: 0.998023

[No partial marking, but full credits are given if its mentioned in percentage or rounded]

5.A.4. What must be the prior probability of cancer, to result in a posterior probability of 0.5 of having cancer given a positive test result?

Answer: 0.048076

[No partial marking, but full credits are given if its mentioned in percentage or rounded]

5.A.5. Now suppose from another study doctors have found out that in addition to the test being positive or negative, the patient's gender is also a factor that contributes to the probability of having cancer. With this knowledge, the doctor needs to formulate the probability the patient has cancer given the test is positive and gender is male (i.e.,

P(Cancer | test positive = p , gender ale = m)). As a probability expert, which of the following statements do you think are true (select all that apply):

Answer: A, C

[Partial credits are given]

We have given full credits for minor rounding error and we have considered small tolerance as well. Some students have given a probability answer as well in percentage(e.g question 1). Full credits have given for that as well. Please be sure to read and follow the instructions in the final exam.

Part B:

Q.5.B.1.1 What is the probability of winning? (0.5 point)

Answer: $0.4 \times \frac{18}{38} + 0.5 \times \frac{19}{38} + 0.1 \times \frac{1}{38} = 0.442105$

Q.5.B.1.2 Suppose that the person loses the round. What is the probability that he bet on red?

Answer: $P(Red|Losttheround) = \frac{P(Losttheround|Red)P(Red)}{P(Losttheround)}$

$= \frac{\frac{19}{38} \times 0.5}{1-0.442105} = 0.448113$

Q.5.B.1.3 If you place a bet on red or black and win, your money is doubled, and if you place your bet on green and win, then your money is tripled. Suppose you bet $1000 in a round of roulette. What is the expected amount that you will have after the round?

Expected amount =
$(1 - 0.442105) \times 0 + 0.4 \times \frac{18}{38} \times 2000 + 0.5 \times \frac{19}{38} \times 2000 + 0.1 \times \frac{1}{38} \times 3000 = 886.842105$

Marks have also been given to those who reported the expected loss which is $1000 - 886.842105

Q.5.B.1.4 What betting strategy should you follow so that on average you do not lose any money while playing the game in Q.5.B.1.3?

Answer: The only way that the expected amount is equal to $1000 is by always betting on red because we have a 0.5 probability of winning by betting on red. In all other cases, the expected amount will be less than $1000.
P(betting on black = 0), P(betting on red) = 1, P(betting on green) = 0

Q.5.B.2.1 Probability of winning starting from $i is equal to the probability of winning the current round times the probability of winning starting from $(i+100) in the next round plus the probability of losing the current round times the probability of winning starting from $(i-100) in the next round.

a = x, b = 1-x

Q.5.B.2.2

$P_i = xP_{i+100} + (1-x)P_{i-100}$

$xP_i + (1-x)P_i = xP_{i+100} + (1-x)P_{i-100}$

$x(P_{i+100} - P_i) = (1-x)(P_i - P_{i-100})$

$P_{i+100} - P_i = \frac{1-x}{x}(P_i - P_{i-100})$

$R = \frac{1-x}{x}$

Using the hint:

$P_{i+100} - P_i = R^{\frac{i}{100}}(P_{100} - P_0) = R^{\frac{i}{100}}P_{100}$

Adding the above equations for i = 0, 1, ... 1900 we get:

$(P_{2000} - P_{1900}) + (P_{1900} - P_{1800}) + (P_{1800} - P_{1700}) + \ldots + (P_{200} - P_{100}) + (P_{100} - P_0) = P_{100}(R^{19} + R^{18} + R^{17} + \ldots + R^1 + R^0)$

$P_{100} = \frac{(P_{2000} - P_0)}{R^0 + R^1 + R^2 + \ldots + R^{19}}$

$P_{i+100} - P_i = R^{\frac{i}{100}}(P_{100} - P_0) = R^{\frac{i}{100}}P_{100}$

Adding the above equations for i = 0, 1, ... 900 we get:

$P_{1000} - P_0 = P_{100}(R^0 + R^1 + R^2 + \ldots + R^9)$

$P_{1000} = \frac{(P_{2000} - P_0)(R^0 + R^1 + R^2 + \ldots + R^9)}{R^0 + R^1 + R^2 + \ldots + R^{19}}$

R = 1 for this part

$P_{1000} = 0.5$

5.B.1.3

Answer: $R = \frac{0.55}{0.45}$

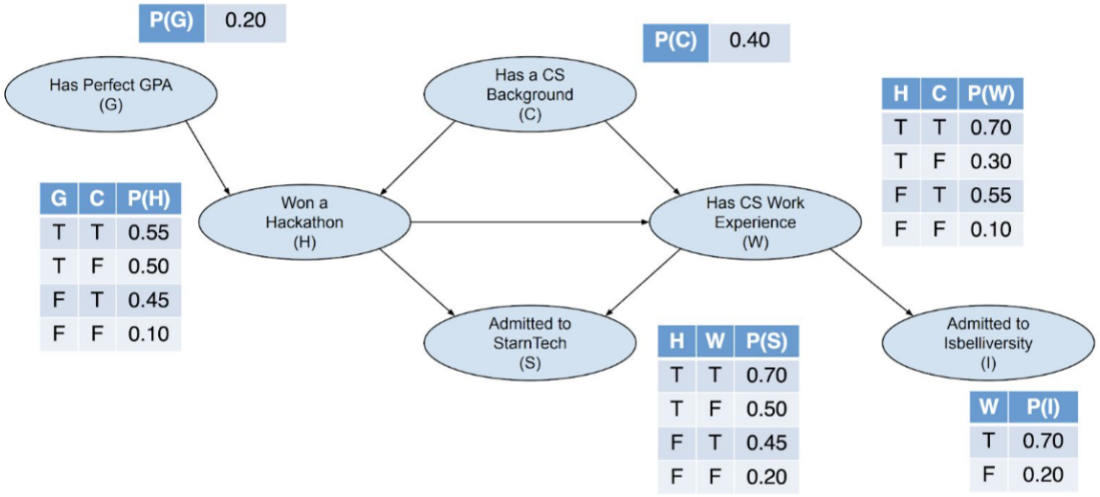
Using the hint for computing the sum $R^0 + R^1 + \ldots + R^n$, we get $P_{1000} = 0.1185005$

Question 6

6. Bayes Nets

(16 points)

StarnTech and Isbelliversity have been in college rivalry with each other over the years. This rivalry is not only in academics and athletics but also in the rigorous admissions process to build a stronger incoming cohort. The Computer Science admissions committee from both universities have recently made public the factors that will be evaluated in order to grant admissions to students in their Master's Program. The following Bayesian Network and the corresponding conditional probability tables describe the admission criteria used.



P(G)	0.20

P(C)	0.40

G	C	P(H)
T	T	0.55
T	F	0.50
F	T	0.45
F	F	0.10

H	C	P(W)
T	T	0.70
T	F	0.30
F	T	0.55
F	F	0.10

H	W	P(S)
T	T	0.70
T	F	0.50
F	T	0.45
F	F	0.20

W	P(I)
T	0.70
F	0.20

Part A

D-Separation and Conceptual Questions (7 points)

Given the above Bayes Network, answer the following questions.

Q.6.A.1 The above Bayes Network is: **(0.5 points)**

- ○ Singly connected
- ◉ Multiply connected

Q.6.A.2 The time and space complexity of Variable Elimination in the worst-case scenario for the above Bayes Net is: **(0.5 points)**

- ○ Linear in the size of the network
- ○ Exponential in the size of the network
- ◉ None of the above

Q.6.A.1 & Q.6.A.2 Explanation:

This Bayes Net is Multiply connected because it is a graph where at least one pair of nodes has more than one path. The complexity of exact inference in Bayesian networks depends strongly on the structure of the network. For multiply connected networks, such as that the one presented in this exam, variable elimination can have exponential time and space complexity in the worst case, even when the number of parents per node is bounded.

However for this particular network it is exponential in the size of the **max parent nodes**, and not in the size of the entire **network**.

Q.6.A.2 (UPDATED RUBRIC)

Linear in the size of the network **(0 points)**

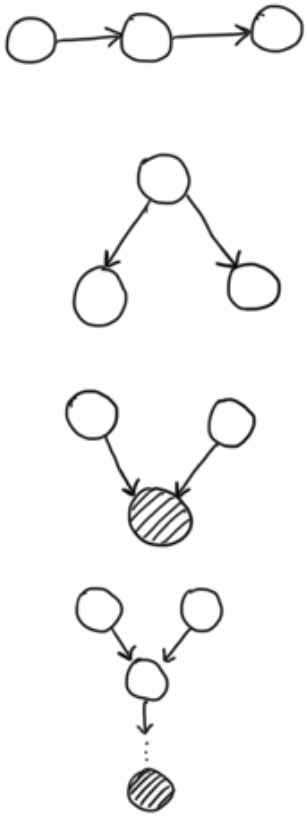
Exponential in the size of the network **(0.25 points)**

None of the above **(0.50 points)**

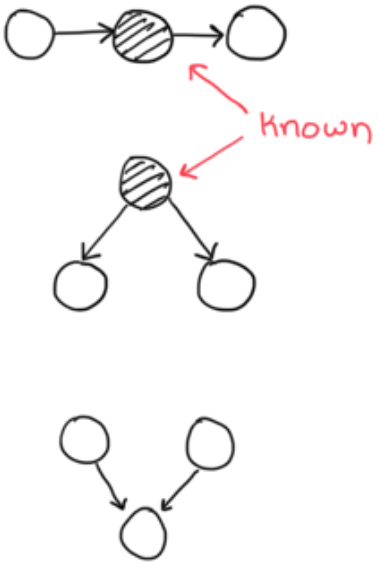
(view updated description above)

For the following questions **Q.6.A.3 - Q.6.A.5** the rules of **D-Separation** as discussed in lecture videos are helpful.

Active Triplets



Inactive Triplets



Q.6.A.3 Consider the claim: "G is conditionally independent of S, given X". For which of the following values of X is the claim true? Choose all such values below. **(1.5 points)**

- [ ] C
- [ ] H
- [ ] W
- [ ] I
- [x] H, W

Q.6.A.4 Consider the claim: "S is conditionally independent of I, given X". For which of the following values of X is the claim true? Choose all such values below. **(1.5 points)**

- [ ] G
- [ ] C
- [ ] H
- [x] W
- [x] H, W

Q.6.A.5 Consider the claim: "G is conditionally independent of I, given X". For which of the following values of X is the claim true? Choose all such values below. **(1.5 points)**

- [ ] C
- [ ] H
- [x] W
- [ ] S
- [ ] C, S

For **Q.6.A.3 - Q.6.A.5** partial credit of **0.3 points** will be given per correct/incorrect option with a total of **1.5 points** per question

Q.6.A.6 Assume **for this question only** that the edge H → W is removed. For the resulting Bayes Network, select all nodes that belong to the Markov Blanket of W: **(1.5 points)**

- [ ] G
- [x] C
- [x] H
- [x] S
- [x] I

Q.6.A.6 Explanation: The Markov Blanket of a node is its Parents, Children, and Children's parents. Therefore, **C, H, S,** and **I** are part of the Markov Blanket for **W**. Partial credit of **0.3 points** given per correct/incorrect option with a total of **1.5 points**.

Q.6.A.6 Assume **for this question only** that the edge H → W is removed. For the resulting Bayes Network, select all nodes that belong to the Markov Blanket of W: **(1.5 points)**

- [ ] G
- [x] C
- [x] H
- [x] S
- [x] I

In this question, given the structure of the Bayesian Network, we know that given W (the only node in the Markov Blanket of I), I is conditionally independent of the remaining of the nodes in the network. Therefore:

$$P(\neg i|\neg w, \neg h, g) = P(\neg i|\neg w) = 1 - P(i|\neg w) = 1 - 0.2 = 0.8$$

Additionally to show that the conditional independence relations hold true and that the rest of the variables can be ignored, consider the following detailed computation which will also reach at the same result:

$$P(i \mid \neg w, \neg h, g) = \alpha \sum_S \sum_C P(i, \neg w, \neg h, g, S, C)$$

$$= \alpha \sum_S \sum_C P(g)\, P(C)\, P(\neg h \mid g, C)\, P(\neg w \mid \neg h, C)\, P(S \mid \neg h, \neg w)\, P(i \mid \neg w)$$

$$= \alpha\, P(g)\, P(i \mid \neg w) \sum_S P(S \mid \neg h, \neg w) \sum_C P(C)\, P(\neg h \mid g, C)\, P(\neg w \mid \neg h, C)$$

$P(\neg i \mid \neg w, \neg h, g)$

$$= \alpha * 0.20 * 0.20 \sum_S P(S \mid \neg h, \neg w) \left[0.40 * 0.45 * 0.45 + 0.6 * 0.5 * 0.9\right]$$

$$= \alpha * 0.20 * 0.20 \sum_S P(S \mid \neg h, \neg w) * 0.351$$

$$= \alpha * 0.20 * 0.20 \left[0.20 * 0.351 + 0.80 * 0.351\right]$$

$$= \alpha * 0.01404$$

$$P(i \mid \neg w, \neg h, g) = \alpha * 0.01404$$

Similarly,

$$P(\neg i \mid \neg w, \neg h, g) = \alpha * 0.20 * 0.80 \left[0.20 * 0.351 + 0.80 * 0.351\right]$$

$$= \alpha * 0.05616$$

$$P(I \mid \neg w, \neg h, g) = \alpha \langle 0.01404, 0.05616 \rangle = \langle 0.2, 0.8 \rangle$$

Full credit is given only if correct answer was calculated, work was showed, and code from Assignment 3 was not used.

Partial credit will be given: **-0.5 points** if input 0.2 instead of 1-0.2=0.8

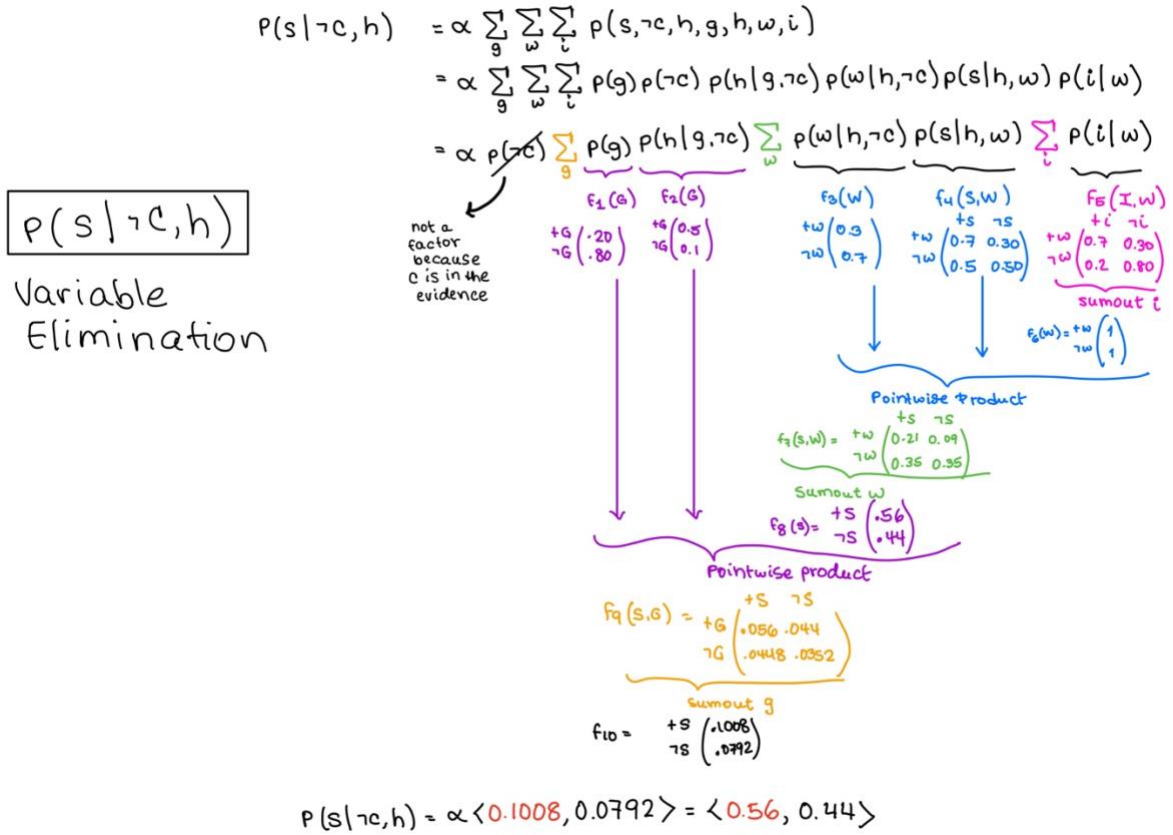
Q.6.B.2.a Compute the probability of Rosa getting admitted to StarnTech given that she does not have a CS background but has won a hackathon. Use Variable Elimination. **(2 points)**

0.56

Q.6.B.2.b Select the variables over which the "sum out" operation is performed in the above question. **(1 point)**

- [x] G
- [ ] C
- [ ] H
- [x] W
- [ ] S
- [x] I

Using Variable Elimination we arrive at the above result:



Q.6.B.2.b. During the process we have to sum out three times, over the variables **G**, **W**, and **I**. The complete version of Variable Elimination or following the pseudocode from Chapter 14 helps to arrive at this answer.

Q.6.B.2.b (UPDATED RUBRIC)

Partial credit will be given **(0.2 points per correct/incorrect choice)**

Correct Answer: **G, W, I**

For **Q.6.B.2.a**, the above computation for Variable Elimination could also be simplified as some terms do not contribute to the final probabilities:

$$P(C = F)\sum_{G} P(G)P(H = T|C = F, G)\sum_{W} P(W|C = F, H = T)P(S|W, H = T)\sum_{I} P(I|W)$$

$$= (0.6)\sum_{G}(0.1\ 0.08)\sum_{W}(0.3\ 0.7)P(W, H = T)$$

$$= (0.6)(0.18)\sum_{W}(0.21\ 0.35\ 0.09\ 0.35)$$

$$= (0.6)(0.18)(0.56\ 0.44)$$

$$= (0.56,\ 0.44)$$

Finally, this could also be solved as showed in the lecture videos by building and reducing tables. For **Q.6.B.2.a** full score is given if arrived at the correct answer, showed reasonable work, and did not use code from Assignment 3.

Q.6.B.3 It is known that the best university is the one that has a lower probability of getting admission.

Q.6.B.3.a Which University is best? **(1 point)**

- ( ) Isbelliversity
- (•) StarnTech
- ( ) Both are the same

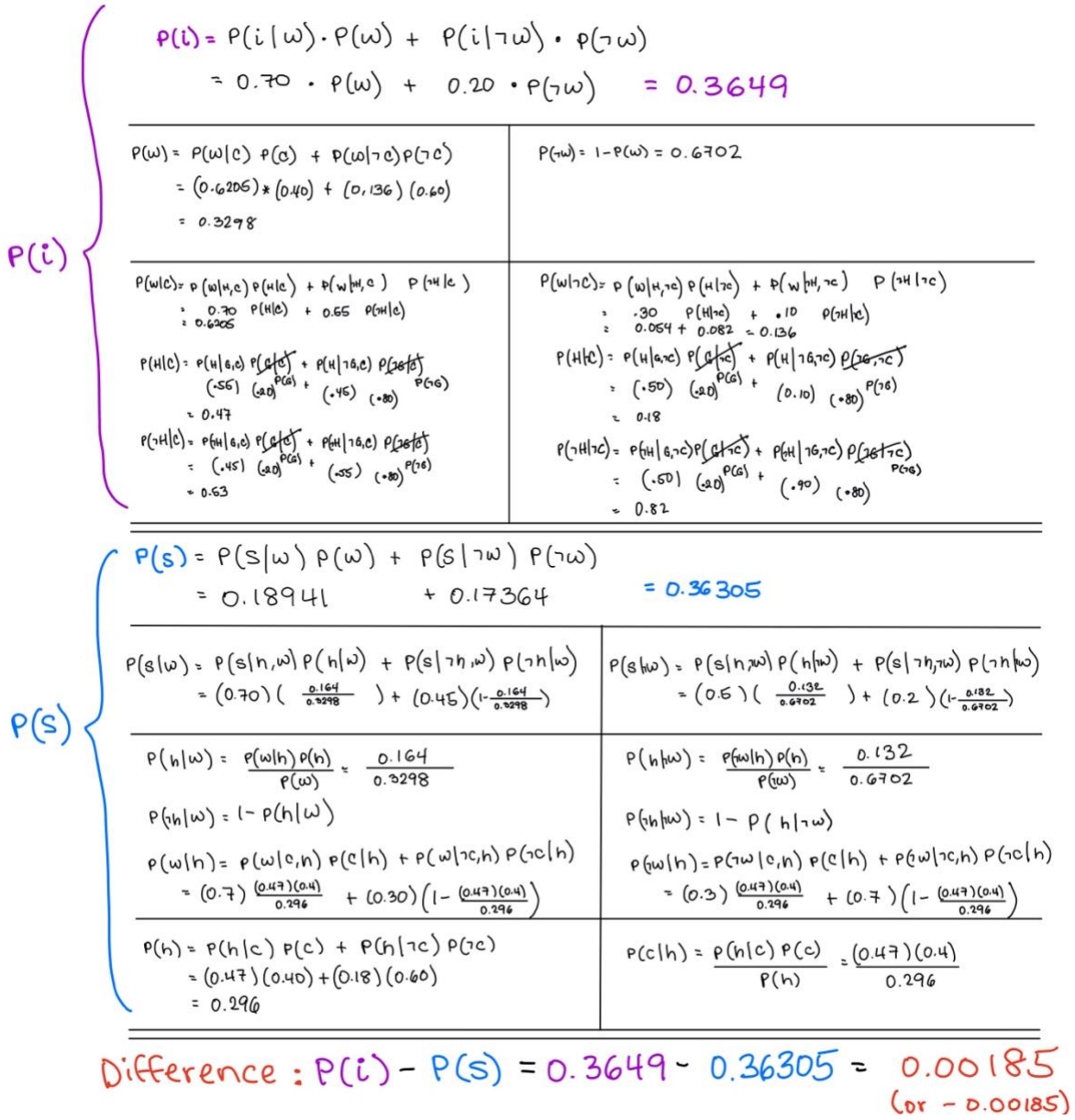
Q.6.B.3.b What is the difference between the two probabilities? **(3 points)**

0.00185

To answer this question two probabilities needed to be calculated and their difference needed to be performed:

$$P(S) - P(I) \text{ or } P(I) - P(S)$$

There are multiple ways to solve for **P(S)** and **P(I)**. One way is to use the law of total probability:



UPDATES TO RUBRIC 03/11/2020 13:09 PM:

Q.6.A.2

Linear in the size of the network **(0 points)**

Exponential in the size of the network **(0.25 points)**

None of the above **(0.50 points)**

(view updated description above)

Q.6.B.2.b

Partial credit will be given **(0.2 points per correct/incorrect choice)**

Correct Answer: **G, W, I**

Q.6.B.3.b

If your values of P(S) and P(I) are correct in the work showed and match the post above, but you subtracted/rounded incorrectly, raise a regrade request for partial credit.

Question 7

Q.7.1.a

State	1	2	3
VB	GP	MK	MS
BB	KK	CR	AA
SV	MK	VB	GP
AA	MK	GP	BV

Q.7.1.b

K	Error
3	3
5	2
7	5

Q.7.1.c

k = 1

Q.7.1.d

False

Q.7.1.e

Supervised learning

Q.7.2.a

Precision = 4/7 (OR) 0.571429

Recall = 4/6 (OR) 0.666667

Q.7.3.a

Iteration 0	[A], [B], [C], [D], [E]
Iteration 1	[A,C],[B],[D],[E]
Iteration 2	[A,C],[B,D],[E]
Iteration 3	[A,C,B,D],[E]
Iteration 4	[A,B,C,D,E]

Q.7.3.b

[A,C,E], [B,D]