Thomas Fiorilla (trf40)

Srikanth Kundeti (sk1799)

Anthony Tiongson (ast119)

Ethan Wang (ew360)

Professor McMahon

CS440 Intro to AI

Assignment 3

April 24, 2020

Question 1

a) Gain (q^+) = \$4000 gain − \$3,000 cost \$1,000 Profit

Gain (q^-) = \$4,000 gain − (\$3,000 cost + \$1,400 repair) = −\$400 Profit

* Since not given a test, mechanic fee is not substituted into cost assessment

b) $P(Pass) = P(Pass \mid q^+)\,P(q^+) + P(Pass \mid q^-)\,P(q^-) = (0.8)(0.7) + (0.35)(0.3) = 0.665$

$P(\neg Pass) = P(fail) = (1 - P(Pass \mid q^+))\,P(q^+) + (1 - P(Pass \mid q^-))\,P(q^-) = (0.2)(0.7) + (0.65)(0.3) = 0.335$

observe $P(Pass) + P(fail) = 0.665 + 0.335 = 1$

$$P(q^+ \mid Pass) = \frac{P(Pass \mid q^+)\,P(q^+)}{P(Pass)} = \frac{(0.8)(0.7)}{(0.665)} = 0.842$$

$$P(q^- \mid Pass) = \frac{P(Pass \mid q^-)\,P(q^-)}{P(Pass)} = \frac{(0.35)(0.3)}{(0.665)} = 0.1578947368$$

$$P(q^+ \mid \neg Pass) = \frac{(1 - P(Pass \mid q^+))\,P(q^+)}{P(\neg Pass)} = \frac{(0.2)(0.7)}{0.335} = 0.4179104478$$

$$P(q^- \mid \neg Pass) = \frac{(1 - P(Pass \mid q^-))\,P(q^-)}{P(\neg Pass)} = \frac{(0.65)(0.3)}{0.335} = 0.5820895522$$

Question 1

(c) It seems the best decision would be the amount of money recieved if the test is passed since there would not need to be repairs done on the car.

$$E[Pass] = P(q^{+} | pass)\, cost(q^{+}) + P(q^{-} | pass)\, cost(q^{-})$$

$$= (\$1000)(0.842) + (-\$400)(0.157)$$

$$= \$779.20$$

$$E[Fail] = P(q^{+} | fail)\, cost(q^{+}) + P(q^{-} | fail)\, cost(q^{-})$$

$$= (\$1000)(0.4179) + (-\$400)(0.582) = \$185.10$$

clearly, the expected utility of passing is greater.

(d) Since \$100 is charged by the mechanic. $cost(q^{+}) = \$1000 - \$100 = \$900$ and $cost(q^{-}) = -\$400 - \$100 = -\$500$

$$E[Pass] = (\$900)(0.842) + (-\$800)(0.157) = \$679.30$$

$$E[Fail] = (\$900)(0.418) + (-\$800)(0.582) = \$85.20$$

Since the expected utility values of both pass and fail given a mechanic is consulted is lower than when he is not consulted, C_1 will not be taken to the mechanic.

Question 2)

(a) Given:

$P(f(left) \mid f(left)) = 0.5$

$P(a(right) \mid f(left)) = 0.2$

$P(a(stay) \mid f(left)) = 0.3$

$P(a(left) \mid f(right)) = 0.2$

$P(a(right) \mid f(right)) = 0.5$

$P(a(stay) \mid f(right)) = 0.3$

$P(a(left) \mid f(same)) = 0.25$

$P(a(right) \mid f(same)) = 0.25$

$P(a(stay) \mid f(same)) = 0.5$

initial

1/5	1/5	1/5	1/5	1/5

0

ant location

$P(f(\text{left of ant})) = \frac{1}{5} = 0.2$

$P(f(\text{right of ant})) = 3 \cdot \frac{1}{5} = \frac{3}{5} = 0.6$

$P(f(\text{same as ant})) = \frac{1}{5} = 0.2$

$P(a(same) \mid f(same))$

or

$$P(a(same)) = P(a(same) \mid f(left)) \cdot P(f(left)) + P(a(same) \mid f(right)) P(f(right)) + P(a(same) \mid f(same)) P(f(same))$$

$$= (0.3)(0.2) + (0.3)(0.6) + (0.5)(0.2) = 0.34$$

$$P(a(R)) = P(a(R) \mid f(L)) P(f(L)) + P(a(R) \mid f(R)) P(f(R)) + P(a(R) \mid f(S)) P(f(S))$$

$$= (0.2)(0.2) + (0.5)(0.6) + (0.25)(0.2) = 0.39$$

$$P(a(L)) = P(a(L) \mid f(L)) P(f(L)) + P(a(L) \mid f(R)) P(f(R)) + P(a(L) \mid f(S)) P(f(S))$$

$$= (0.5)(0.2) + (0.2)(0.6) + (0.25)(0.2) = 0.27$$

*Notice $P(a(S)) + P(a(R)) + P(a(L)) = 0.34 + 0.39 + 0.27 = 1.0$

0.2	0.2	0.2	0.2	0.2

$$P(f(C_1) \mid a(R)) = \frac{P(a(R) \mid f(L))\, P(f(L))}{P(a(R))} = \frac{(0.2)(0.2)}{(0.39)} = 0.1025641026$$

$$P(f(C_2) \mid a(R)) = \frac{P(a(R) \mid f(S))\, P(f(S))}{P(a(R))} = \frac{(0.25)(0.2)}{(0.39)} = 0.1282051282$$

$$P(f(C_3) \mid a(R)) = P(f(C_4) \mid a(R)) = P(f(C_5) \mid a(R)) = \frac{1}{3} \cdot P(f(R) \mid a(R)) = \frac{1}{3}\left[\frac{P(a(R) \mid f(R))\, P(f(R))}{P(a(R))}\right]$$

$$= \frac{1}{3}\, \frac{(0.5)(0.6)}{(0.39)} = 0.2564102564$$

$$P(f(C_1) \mid a(R)) + P(f(C_2) \mid a(R)) + P(f(C_3) \mid a(R)) + P(f(C_4) \mid a(R)) + P(f(C_5) \mid a(R))$$

$$= 0.1025641026 + 0.1282051282 + 3(0.2564102564) = 1$$

0.1025641026	0.1282051282	0.2564102564	0.2564102564	0.2564102564

ant location after 1 step right.

$$P(f(\text{left})) = 0.1025641026 + 0.1282051282 = 0.2307692308$$

$$P(f(\text{same})) = 0.2564102564$$

$$P(f(\text{right})) = 0.5128205128$$

$$P(a(\text{same})) = P(a(S) \mid f(L))\, P(f(L)) + P(a(S) \mid f(R))\, P(f(R)) + P(a(S) \mid f(S))\, P(f(S))$$

$$= (0.3)(0.2307692308) + (0.3)(0.5128205128) + (0.5)(0.2564102564)$$

$$= 0.3512820513$$

$$P(a(R)) = P(a(R) \mid f(L))\, P(f(L)) + P(a(R) \mid f(R))\, P(f(R)) + P(a(R) \mid f(S))\, P(f(S))$$

$$= (0.2)(0.2307692308) + (0.5)(0.5128205128) + (0.25)(0.2564102564)$$

$$= 0.3666666667$$

$$P(a(L)) = P(a(L) \mid f(L))\, P(f(L)) + P(a(L) \mid f(R))\, P(f(R)) + P(a(L) \mid f(S))\, P(f(S))$$

$$= (0.5)(0.2307692308) + (0.2)(0.5128205128) + (0.2564102564)(0.25)$$

$$= 0.2820512821 \qquad P(a(S)) + P(a(R)) + P(a(L)) = 1 \quad (\text{check})$$

0.1025641026	0.1282051282	0.2564102564	0.2564102564	0.2564102564

1 current location 2

$$P(f(c_1)\mid a(R)) = P(f(c_2)\mid a(R)) = \frac{1}{2}\left[\frac{P(a(R)\mid f(L))\cdot P(f(L))}{P(a(R))}\right] = \frac{(0.2)\ (0.2307692308)}{0.3666666667\,(2)}$$

$$= 0.0629370629$$

$$P(f(c_3)\mid a(R)) = P(f(\text{same})\mid a(\text{Right})) = \frac{P(a(R)\mid f(S))\,P(f(S))}{P(a(R))} = \frac{(0.25)(0.2564102564)}{(0.3666666667)}$$

$$= 0.1748251748$$

$$P(f(c_4)\mid a(R)) = P(f(c_5)\mid a(R)) = \frac{1}{2}\,P(f(R)\mid a(R)) = \frac{P(a(R)\mid f(R))\,P(f(R))}{2\,(P(a(R)))} = \frac{(0.5)(0.5128205128)}{2(0.3666666667)}$$

$$= 0.3496503496$$

$$\sum_{i}^{5} P(f(c_i)\mid a(R)) = 2(0.0629370629) + 0.1748251748 + 2(0.3496503496) = 0.9999999998 \approx 1$$

0.0629370629	0.0629370629	0.1748251748	0.3496503496	0.3496503496

3

1 2

$P(f(L)) = \cancel{\ldots}\ 0.3006993006$

$P(f(R)) = \cancel{\ldots}\ 0.3496503496$

$P(f(S)) = \cancel{\ldots}\ 0.3496503496$

$$P(a(S)) = P(a(S)\mid f(L))\,P(f(L)) + P(a(S)\mid f(R))\,P(f(R)) + P(a(S)\mid f(S))\,P(f(S))$$

$$= (0.3)\,(\overset{0.3006993006}{\cancel{\ldots}}) + (0.3)\,(\overset{0.3496503496}{\cancel{\ldots}}) + (0.5)\,(\overset{0.3496503496}{\cancel{\ldots}})$$

$$= \cancel{0.3349650349}\ 0.3699301431$$

$$P(a(R)) = P(a(R)\mid f(L))\,P(f(L)) + P(a(R)\mid f(R))\,P(f(R)) + P(a(R)\mid f(S))\,P(f(S))$$

$$= (0.2)\,(\overset{0.3006993006}{\cancel{\ldots}}) + (0.5)\,(\overset{0.3496503496}{\cancel{\ldots}}) + (0.25)\,(\overset{0.3496503496}{\cancel{\ldots}})$$

$$= \cancel{\ldots}\ 0.3223776223$$

$$P(a(L)) = P(a(L)\mid f(L))\,P(f(L)) + P(a(L)\mid f(R))\,P(f(R)) + P(a(L)\mid f(S))\,P(f(S))$$

$$= (0.5)\,(\overset{0.3006993006}{\cancel{\ldots}}) + (0.2)\,(\overset{0.3496503496}{\cancel{\ldots}}) + (0.25)\,(\overset{0.3496503496}{\cancel{\ldots}})$$

$$= \cancel{\ldots}\ 0.3076923076$$

$$P(a(S)) + P(a(R)) + P(a(L)) = \overset{1.000000073}{\cancel{\ldots}} \approx 1$$

0.0629370629	0.0629370629	0.1748251748	0.3496503496	0.3496503496

o
ant location
after 1st step
right,
next step left

$$P(f(C_1) \vee f(C_2) \vee f(C_3) \mid a(L)) = \frac{1}{3} P(f(L) \mid a(L)) = \frac{P(a(L) \mid f(L)) \cdot P(f(L))}{3 \cdot P(a(L))} = \frac{(0.5)(0.3006993006)}{3(0.3076923076)}$$

$$= 0.4886363636\left(\frac{1}{3}\right) = 0.1628787879$$

$$P(f(C_4) \mid a(L)) = P(f(S) \mid a(L)) = \frac{P(a(L) \mid f(S))\, P(f(S))}{P(a(L))} = \frac{(0.25)(0.3496503496)}{(0.3076923076)}$$

$$= 0.2840909091$$

$$P(f(C_5) \mid a(L)) = P(f(R) \mid a(L)) = \frac{P(a(L) \mid f(R))\, P(f(R))}{P(a(L))} = \frac{(0.2)(\cancel{0.6993006992})\ 0.349503496}{(\cancel{0.2465034964})\ 0.3076923076}$$

$$= 0.2276325375$$

$$\sum_{i=1}^{5} P(f(C_i) \mid a(L)) = 1.00035981 \approx 1$$

0.1628787879	0.1628787829	0.1628787829	0.2840909091	0.2276325375

o
current
location

$P(f(L)) = 0.3257575758$
$P(f(S)) = 0.1628787879$
$P(f(R)) = 0.5681818182$

(b) $P(a(S)) = (0.3)(0.325) + (0.3)(0.163) + (0.5)(0.568) = 0.4304$

$P(a(R)) = (0.2)(0.325) + (0.5)(0.163) + (0.25)(0.568) = 0.2885$

$P(a(L)) = (0.5)(0.325) + (0.2)(0.163) + (0.25)(0.568) = 0.3371$

1.056
≈ 1
due to rounding.

Question 3 [25 points]:

a) Prove that

$$P(X \mid MB(X)) = \alpha P(X \mid U_1, \ldots, U_m) \prod_{Y_i}^{n} P(Y_i \mid Z_{i1}, \ldots)$$

where $MB(X)$ is the Markov Blanket of variable X.

The probability of the Markov Blanket of variable X can be written as:

$$P(U_1, \ldots, U_m, X, Y_1, \ldots, Y_n, Z_{11}, \ldots, Z_{nj})$$

$$= P(U_1 \cap \ldots \cap U_m \cap X \cap Y_1 \cap \ldots \cap Y_n \cap Z_{11} \cap \ldots \cap Z_{nj})$$

$$= [\prod_{i=1}^{m} P(U_i)] P(X \mid U_1, \ldots, U_m) [\prod_{i=1}^{n} P(Y_i \mid X, Z_{i1}, \ldots, Z_{ij}) \prod_{k=1}^{j} P(Z_{ik})]$$

The Markov Blanket of variable X is the set:

$$MB(X) = \{U_1, \ldots, U_i, Y_1, \ldots, Y_n, Z_{11}, \ldots, Z_{nj}\}$$

The conditional probability of X given $MB(X)$ can be written as:

$$P(x \mid MB(x)) = \frac{P(x \cap MB(x))}{P(MB(x))}$$

The numerator $P(x \cap MB(x))$ can be equivalently expressed as the probability of the Markov Blanket of variable x:

$$[\prod_{i=1}^{m} P(u_i)] P(x \mid u_1, \ldots, u_m) [\prod_{i=1}^{n} P(y_i \mid x, z_{i1}, \ldots, z_{ij}) \prod_{k=1}^{j} P(z_{ik})]$$

The denominator $P(MB(x))$ can be written as:

$$P(u_1, \ldots, u_m, y_1, \ldots, y_n, z_{11}, \ldots, z_{nj})$$

$$= P(u_1 \cap \ldots \cap u_m \cap y_1 \cap \ldots \cap y_n \cap z_{11} \cap \ldots \cap z_{nj})$$

$$= [\prod_{i=1}^{m} P(u_i)][\sum_{x} \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij}) \prod_{k=1}^{j} P(z_{ik})]$$

Consequently, we have:

$$P(x \,|\, MB(x)) = \frac{[\prod_{i=1}^{m} P(u_i)] P(x \,|\, u_1, \ldots, u_m) [\prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij}) \prod_{k=1}^{j} P(z_{ik})]}{[\prod_{i=1}^{m} P(u_i)][\sum_{x} \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij}) \prod_{k=1}^{j} P(z_{ik})]}$$

$$= \frac{[\prod_{i=1}^{m} P(u_i)] P(x \,|\, u_1, \ldots, u_m) [\prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})][\prod_{i=1}^{n} \prod_{k=1}^{j} P(z_{ik})]}{[\prod_{i=1}^{m} P(u_i)][\sum_{x} \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})][\prod_{i=1}^{n} \prod_{k=1}^{j} P(z_{ik})]}$$

$$= \frac{P(x \,|\, u_1, \ldots, u_m) \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})}{\sum_{x} \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})}$$

The ratio $\dfrac{1}{\sum_{x} \prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})}$ represents the normalization constant α.

Thus, we have proven:

$$P(x \,|\, MB(x)) = \alpha P(x \,|\, u_1, \ldots, u_m)[\prod_{i=1}^{n} P(y_i \,|\, x, z_{i1}, \ldots, z_{ij})]$$

$Q.E.D.$

b) Consider the query

$$P(Rain \mid Sprinkler = true, WetGrass = true)$$

in the Rain/Sprinkler network and how MCMC would answer it. How many possible states are there for the approach to consider given the network and the available evidence variables?

Considering the number of evidence variables is 2, $Sprinkler = true$ and $WetGrass = true$, and the nonevidence variables for the query are $Rain$ and $Cloudy$ since the Markov Blanket for node $Rain$ is the set $\{Sprinkler, WetGrass, Cloudy\}$, the initial state for the network can be represented as $[true, true, R_i, C_i]$ where R_i and C_i can either be randomly initialized as $true$ or $false$. This gives us 4 possible states for the initial approach to consider: $[true, true, false, false]$, $[true, true, false, true]$, $[true, true, true, false]$, and $[true, true, true, true]$.

c) Using the query in 3b, calculate the transition matrix Q that stores the probabilities $P(y \to y')$ for all the states y, y'. If the Markov Chain has n states, then the transition matrix has size $n \times n$ and you should compute n^2 probabilities.

For the query $P(Rain \mid Sprinkler = true, WetGrass = true)$, we know the Markov Chain has 4 states so the resulting transition matrix Q will be a 4×4 matrix with 16 probabilities for $y \to y'$ where y represents a "current" state for the nonevidence variables $Rain$ and $Cloudy$ and y' is a possible next state that it can transition to. Of the 16 probabilities, we know that 4 will equal 0 for when the following state has both nonevidence variables change, like from

$Rain, Cloudy \rightarrow \neg Rain, \neg Cloudy$, since in Markov Chain Monte Carlo only a single variable is allowed to change at each transition. The matrix Q can then be initially visualized as:

$$
\begin{array}{c|c}
 & \begin{array}{cccc} Rain, Cloudy & Rain, \neg Cloudy & \neg Rain, Cloudy & \neg Rain, \neg Cloudy \end{array} \\
\begin{array}{c} Rain, Cloudy \\ Rain, \neg Cloudy \\ \neg Rain, Cloudy \\ \neg Rain, \neg Cloudy \end{array} &
\begin{bmatrix} X & X & X & 0 \\ X & X & 0 & X \\ X & 0 & X & X \\ 0 & X & X & X \end{bmatrix}
\end{array}
$$

To calculate the transition probabilities, we first calculate the the MCMC Gibbs sampling probabilities for $Rain$ and $Cloudy$ conditioned on the evidence variables $Sprinkler$ and $WetGrass$ in the respective Markov Blankets of $Rain$ and $Cloudy$:

$$P(Rain \mid Cloudy, Sprinkler, WetGrass) = \alpha P(Rain \mid Cloudy)P(WetGrass \mid Rain, Sprinkler)$$

$$= \alpha(0.8)(0.99) = \frac{0.792}{0.972} = \frac{22}{27}$$

$$P(\neg Rain \mid Cloudy, Sprinkler, WetGrass) = \alpha P(\neg Rain \mid Cloudy)P(WetGrass \mid \neg Rain, Sprinkler)$$

$$= \alpha(0.2)(0.9) = \frac{0.18}{0.972} = \frac{5}{27}$$

$$P(Rain \mid \neg Cloudy, Sprinkler, WetGrass) = \alpha P(Rain \mid \neg Cloudy)P(WetGrass \mid Rain, Sprinkler)$$

$$= \alpha(0.21)(0.99) = \frac{0.198}{0.918} = \frac{11}{51}$$

$$P(\neg Rain \mid \neg Cloudy, Sprinkler, WetGrass) = \alpha P(\neg Rain \mid \neg Cloudy) P(WetGrass \mid \neg Rain, Sprinkler)$$

$$= \alpha(0.8)(0.9) = \frac{0.72}{0.918} = \frac{40}{51}$$

$$P(Cloudy \mid Rain, Sprinkler) = \alpha P(Cloudy) P(Rain \mid Cloudy) P(Sprinkler \mid Cloudy)$$

$$= \alpha(0.5)(0.8)(0.1) = \frac{0.04}{0.09} = \frac{4}{9}$$

$$P(\neg Cloudy \mid Rain, Sprinkler) = \alpha P(\neg Cloudy) P(Rain \mid \neg Cloudy) P(Sprinkler \mid \neg Cloudy)$$

$$= \alpha(0.5)(0.2)(0.5) = \frac{0.05}{0.09} = \frac{5}{9}$$

$$P(Cloudy \mid \neg Rain, Sprinkler) = \alpha P(Cloudy) P(\neg Rain \mid Cloudy) P(Sprinkler \mid Cloudy)$$

$$= \alpha(0.5)(0.2)(0.1) = \frac{0.01}{0.21} = \frac{1}{21}$$

$$P(\neg Cloudy \mid \neg Rain, Sprinkler) = \alpha P(\neg Cloudy) P(\neg Rain \mid \neg Cloudy) P(Sprinkler \mid \neg Cloudy)$$

$$= \alpha(0.5)(0.8)(0.5) = \frac{0.20}{0.21} = \frac{20}{21}$$

We can now use these values to compute each transition probability. For transitions where the transition state equals the initial state, both samplings of those variables are taken into account in the form $\frac{1}{2}[SamProb_1 + SamProb_2]$.

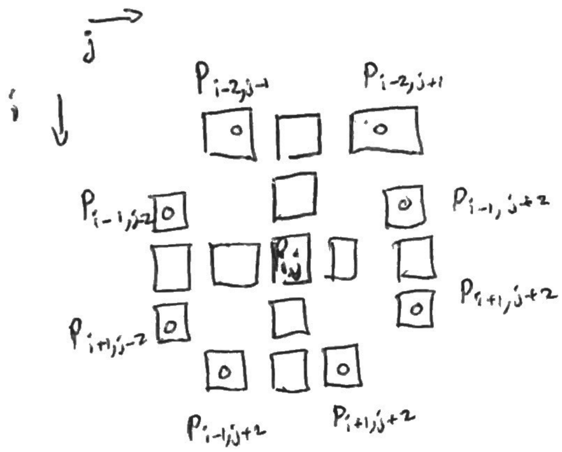
For transitions where only one variable changes, only the sampling probability of the resulting variable is considered in the form $\frac{1}{2}SamProb$. The matrix Q is then:

$$
\begin{array}{c|cccc}
 & Rain, Cloudy & Rain, \neg Cloudy & \neg Rain, Cloudy & \neg Rain, \neg Cloudy \\
\hline
\begin{array}{c} Rain, Cloudy \\ Rain, \neg Cloudy \\ \neg Rain, Cloudy \\ \neg Rain, \neg Cloudy \end{array} &
\multicolumn{4}{c}{\begin{bmatrix}
\frac{17}{27} & \frac{5}{18} & \frac{5}{54} & 0 \\
\frac{2}{9} & \frac{59}{153} & 0 & \frac{20}{51} \\
\frac{11}{27} & 0 & \frac{22}{189} & \frac{10}{21} \\
0 & \frac{11}{102} & \frac{1}{42} & \frac{310}{357}
\end{bmatrix}}
\end{array}
$$

Question 4 Introduction to Artificial Intelligence [redacted]

(a) The variables are boolean, stating whether a pawn / knight is placed on the board or not. There is a variable for every grid on the board. The possible values are P or $\neg P$ (1 or 0) where P (or 1) is a pawn that is placed, and $\neg P$ (or 0) is a pawn that cannot be placed on the board, so it must be left empty. The location of a variable is given by the subscript under the variable P : $P_{i,j}$ for a matrix (ith row, jth column).

(b)



Constraints:

$$P_{i,j} \rightarrow \neg P_{i-2,j-1} \wedge \neg P_{i-2,j+1} \wedge \neg P_{i-1,j-2} \wedge \neg P_{i+1,j-2} \wedge \neg P_{i-1,j+2} \wedge \neg P_{i+1,j+2} \wedge \neg P_{i-1,j+2} \wedge \neg P_{i+1,j+2}$$

$$\neg P_{i,j} \rightarrow P_{i-2,j-1} \vee P_{i-2,j+1} \vee P_{i-1,j-2} \vee P_{i+1,j-2} \vee P_{i-1,j+2} \vee P_{i+1,j+2} \vee P_{i+1,j+2} \vee P_{i-1,j+2}$$

(c) The CSF algorithm (Backtracking Algorithm) can be used to place the pawns in such a way that there would be no conflict, given the constraints, with the arrangement.
Much like the n-queens problem, to solve this problem it's best to place many random pieces on the board until all constraints are in conflict, then backtrack or simply backtrack after initially placing all the pawns on the board. When deciding where to place a pawn, we may use the Min-Conflict algorithm variation (page 225) to calculate the heuristic of a pawn placement based on the number of conflicts that arise as a result for any single placement. For example, a pawn placement where a conflict of 0 versus a conflict of 3 (where 3 constraints are violated) arises, the clear choice would be the placement on the grid with a heuristic of zero.