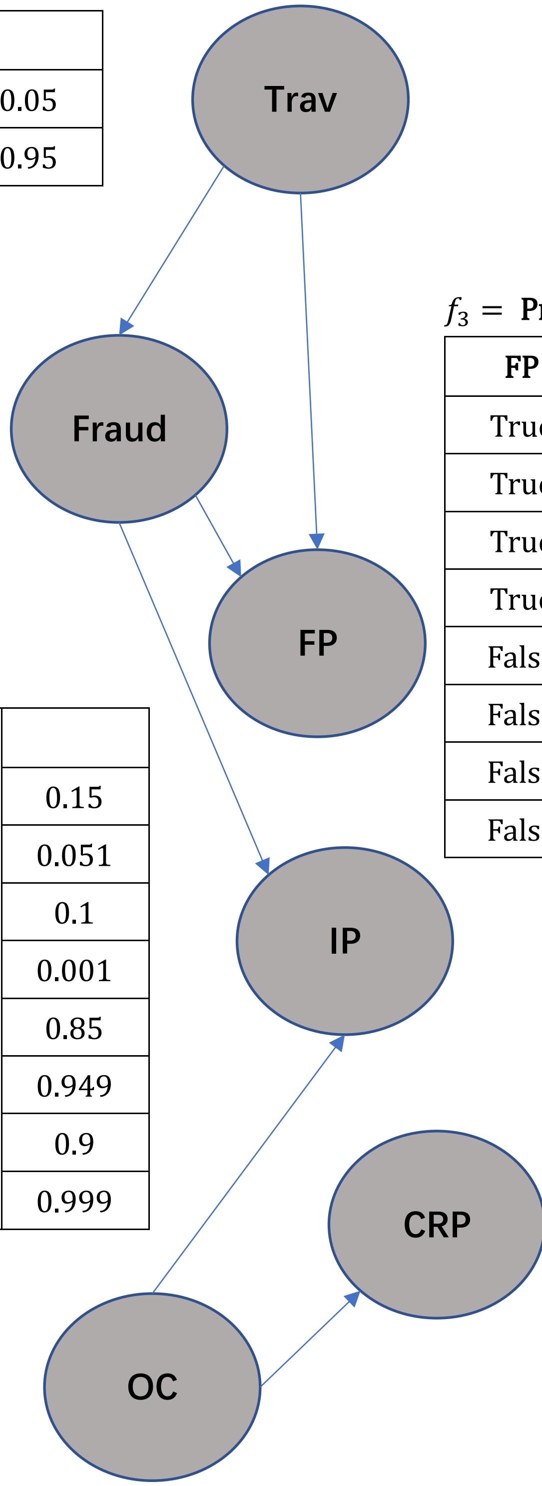
CS486 - Homework Assignment 2 Chen Lang 20528863

2 (a)

$f_1 =$ **Pr(Trav)**

Trav	
True	0.05
False	0.95



$f_2 =$ **Pr(Fraud | Trav)**

Fraud	**Trav**	
True	True	0.01
True	False	0.004
False	True	0.99
False	False	0.996

$f_3 =$ **Pr(FP | Fraud, Trav)**

FP	**Fraud**	**Trav**	
True	True	True	0.9
True	True	False	0.1
True	False	True	0.9
True	False	False	0.01
False	True	True	0.1
False	True	False	0.9
False	False	True	0.1
False	False	False	0.99

$f_4 =$ **Pr(IP | Fraud, OC)**

IP	**Fraud**	**OC**	
True	True	True	0.15
True	True	False	0.051
True	False	True	0.1
True	False	False	0.001
False	True	True	0.85
False	True	False	0.949
False	False	True	0.9
False	False	False	0.999

$f_5 =$ **Pr(CRP | OC)**

CRP	**OC**	
True	True	0.1
True	False	0.01
False	True	0.9
False	False	0.99

$f_6 =$ **Pr(OC)**

OC	
True	0.8
False	0.2

2 (b) (1) Pr(Fraud)

Elimination procedure:

Multiply f_1 and $f_2 = f_7$

Fraud	**Trav**	
True	True	0.0005
True	False	0.0038
False	True	0.0495
False	False	0.9462

Sumout **Trav** in f_7 to f_8

Fraud	
True	0.0043
False	0.9957

$\Pr(\text{Fraud}) = 0.0043 = 0.43\%$

(b) (2) $\Pr(\text{Fraud}|\text{FP} = \text{true}, \text{IP} = \text{false}, \text{CRP} = \text{true})$

Elimination procedure:

Restrict **FP** in f_3 to f_7

Fraud	**Trav**	
True	True	0.9
True	False	0.1
False	True	0.9
False	False	0.01

Restrict **IP** in f_4 to f_8

Fraud	**OC**	
True	True	0.85
True	False	0.949
False	True	0.9
False	False	0.999

Restrict **CRP** in f_5 to f_9

OC	
True	0.1
False	0.01

Multiply all factors with **Trav** to f_{10}

Fraud	**Trav**	
True	True	0.00045
True	False	0.00038
False	True	0.04455
False	False	0.009462

Sumout **Trav** in f_{10} to f_{11}

Fraud	
True	0.00083
False	0.054012

Multiply all factors with **OC** to f_{12}

Fraud	**OC**	
True	True	0.068
True	False	0.001898
False	True	0.072
False	False	0.001998

Sumout **OC** in f_{12} to f_{13}

Fraud	
True	0.069898
False	0.073998

Multiply the rest factors:

Fraud	
True	$5.8 * 10^{-5}$
False	0.004

After normalizing:

Fraud	
True	0.0143
False	0.9857

$$\Pr(\text{Fraud}|\text{FP} = \text{true}, \text{IP} = \text{false}, \text{CRP} = \text{true}) = 0.0143 = 1.43\%$$

2 (c) $\Pr(\text{Fraud}|\text{FP} = \text{true}, \text{IP} = \text{false}, \text{CRP} = \text{true}, \text{Trav} = \text{true})$

Elimination procedure:

Restrict **FP** in f_3 to f_7

Fraud	**Trav**	
True	True	0.9
True	False	0.1
False	True	0.9
False	False	0.01

Restrict **IP** in f_4 to f_8

Fraud	**OC**	
True	True	0.85
True	False	0.949
False	True	0.9
False	False	0.999

Restrict **CRP** in f_5 to f_9

OC	
True	0.1
False	0.01

Restrict **Trav** in f_1 to f_{10}

Trav	
True	0.05

Restrict **Trav** in f_2 to f_{11}

Fraud	
True	0.01
False	0.99

Restrict **Trav** in f_7 to f_{12}

Fraud	
True	0.9
False	0.9

Multiply all factors with **OC** to f_{13}

Fraud	**OC**	
True	True	0.068
True	False	0.001898
False	True	0.072
False	False	0.001998

Sumout **OC** in f_{13} to f_{14}

Fraud	
True	0.069898
False	0.073998

Multiply the rest factors:

Fraud	
True	$3.1 * 10^{-5}$
False	0.0033

After normalizing:

Fraud	
True	0.00945
False	0.99055

$\Pr(\text{Fraud}|\text{FP} = \text{true}, \text{IP} = \text{false}, \text{CRP} = \text{true}, \text{Trav} = \text{true}) = 0.00945 = 0.945\%$

2 (d)

The action taken is making an unimportant purchase over internet first, which makes $\text{CRP} = \text{true}$.

Doing so, the system will generate:

$$\Pr(\text{Fraud}|\text{IP} = \text{true})$$

Restrict **IP** in f_4 to f_7

Fraud	**OC**	
True	True	0.15
True	False	0.051
False	True	0.1
False	False	0.001

Multiply all factors with **Trav** to f_8

FP	**Fraud**	**Trav**	
True	True	True	0.00045
True	True	False	0.00038
True	False	True	0.04455
True	False	False	0.009462
False	True	True	$5 * 10^{-5}$
False	True	False	0.00342
False	False	True	0.00495
False	False	False	0.936738

Sumout **Trav** in f_8 to f_9

FP	**Fraud**	
True	True	0.00083
True	False	0.054012
False	True	0.00347
False	False	0.941688

Multiply all factors with **FP** to f_{10}

FP	**Fraud**	
True	True	0.00083
True	False	0.054012
False	True	0.00347
False	False	0.941688

Sumout **FP** in f_{10} to f_{11}

Fraud	
True	0.0043
False	0.9957

Multiply all factors with **OC** to f_{12}

CRP	**Fraud**	**OC**	
True	True	True	0.012
True	True	False	0.000102
True	False	True	0.008
True	False	False	$2 * 10^{-6}$
False	True	True	0.108
False	True	False	0.010098
False	False	True	0.072
False	False	False	0.000198

Sumout **OC** in f_{12} to f_{13}

CRP	**Fraud**	
True	True	0.012102
True	False	0.008002
False	True	0.118098
False	False	0.072198

Multiply all factors with **CRP** to f_{14}

CRP	**Fraud**	
True	True	0.012102
True	False	0.008002
False	True	0.118098
False	False	0.072198

Sumout **CRP** in f_{14} to f_{15}

Fraud	
True	0.1302
False	0.0802

Multiply the rest factors:

Fraud	
True	0.00056
False	0.07986

After normalizing:

Fraud	
True	0.00696213393
False	0.99303786607

$$\Pr(\text{Fraud}|\text{IP} = \text{true}) = 0.00696213393 = 0.696213393\%$$

After that, the system will generate:

$$\Pr(\text{Fraud}|\text{IP} = \text{true}, \text{CRP} = \text{true})$$

Restrict **IP** in f_4 to f_7

Fraud	OC	
True	True	0.15
True	False	0.051
False	True	0.1
False	False	0.001

Restrict **CRP** in f_5 to f_8

OC	
True	0.1
False	0.01

Multiply all factors with **Trav** to f_9

FP	Fraud	Trav	
True	True	True	0.00045
True	True	False	0.00038
True	False	True	0.04455
True	False	False	0.009462
False	True	True	$5 * 10^{-5}$
False	True	False	0.00342
False	False	True	0.00495
False	False	False	0.936738

Sumout **Trav** in f_9 to f_{10}

FP	**Fraud**	
True	True	0.00083
True	False	0.054012
False	True	0.00347
False	False	0.941688

Multiply all factors with **FP** to f_{11}

FP	**Fraud**	
True	True	0.00083
True	False	0.054012
False	True	0.00347
False	False	0.941688

Sumout **FP** in f_{11} to f_{12}

Fraud	
True	0.0043
False	0.9957

Multiply all factors with **OC** to f_{13}

CRP	**Fraud**	
True	True	0.012
True	False	0.000102
False	True	0.008
False	False	$2 * 10^{-6}$

Sumout **FP** in f_{13} to f_{14}

Fraud	
True	0.012102
False	0.008002

Multiply the rest factors:

Fraud	
True	$5.2 * 10^{-6}$
False	0.079676

After normalizing:

Fraud	
True	0.00648890285
False	0.99351109715

$\Pr(\text{Fraud}|\text{IP} = \text{true}, \text{CRP} = \text{true}) = 0.00648890285 = 0.648890285\%$

Above all,

$\Pr(\text{Fraud}|\text{IP} = \text{true}) - \Pr(\text{Fraud}|\text{IP} = \text{true}, \text{CRP} = \text{true})$

$= 0.696213393\% - 0.648890285\%$

$= 0.047323108\%$

The probability of a fraud gets reduced is 0.047323108%

3 (a)

(i) No, D and G are dependent. There is a path from D to G which is not blocked since there is no evidence in that path. So, D and G are dependent.

(ii) No, D and G are dependent. There is an evidence F in one path from D to G, but there is another path from D to G that is not blocked. So, D and G are dependent.

(iii) Yes, A and G are independent. The indirect path from A to A and other indirect paths from G to C enter B, but there is no evidence for B, and B does not have descendent. So, B blocks all paths between A and G, and they are d-separated. So, A and G are independent.

(iv) No, A and G are dependent. The indirect path from G to D goes into C, and the indirect path from A to B goes out C. C is not in the evidence set, so this path between A and G is not blocked. So, A and G are dependent.

(v) Yes, A and G are independent. Firstly, the indirect path from A to B leaves C, and indirect path from G to D enters C. Also, there is an evidence for C. So, this path is blocked. Secondly, the indirect path from A to B and the indirect path from G to E both leave C, and there is an evidence for C. So, this path is also blocked. Thus, paths between A and G are blocked, and they are d-separated. So, A and G are independent.

(vi) Yes, A and G are independent. Firstly, the indirect path from A to C and the indirect path from G to F both leave D, and there is an evidence for D. So, this path is blocked. Secondly, the indirect path from A to C and the indirect path from G to F both enter E, and there is no evidence for E. So, this path is also blocked. Thus, paths between A and G are blocked, and they are d-separated. So, A and G are independent.

(vii) No, A and G are dependent. The indirect path from G to F and the indirect path from A to C both enter E. E is in the evidence set, so this path between A and G is not blocked. So, A and G are dependent.

3 (b)

(1) C is relevant since C is the query variable.

(2) D is relevant since D is the parent of C, and C is relevant.

(3) E is relevant since E is in evidence set, and E is the descendent of a relevant node C.

(4) F is relevant since F is the parent of E, and F is relevant.

Above all, the subset of relevant variables that is sufficient to answer this query is {C, D, E, F}.