1

En una planta nuclear hay una alarma que se activa cuando un indicador de temperatura excede un umbral. El indicador mide la temperatura del núcleo. Considera las variables booleanas A (alarma), D_A (alarma defectuosa), D_I (indicador defectuoso) y las variables enteras I (lectura del indicador) y T (temperatura real del núcleo).

- Dibuja 3 redes bayesianas válidas diferentes que capturen el comportamiento del proceso, entre ellas incluye aquella que capture el mayor número de independencias condicionales con el menor número de arcos

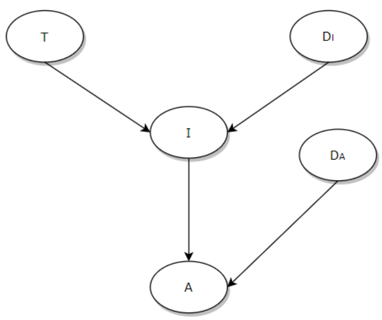


Figure 1: Red bayesiana 1

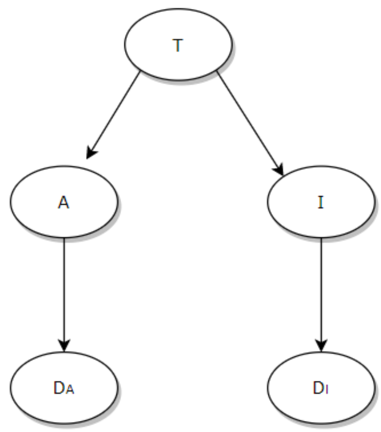


Figure 2: Red bayesiana 2

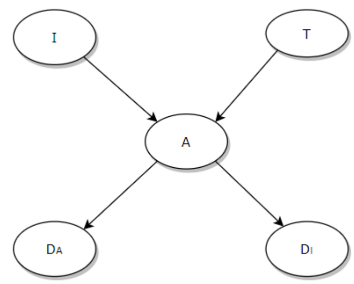


Figure 3: Red bayesiana 3

- Discute el modelo representado por cada una de las redes.
  Por como lo entendi y un poco de sentido común, si la temperatura aumenta, pueden ocurrir mas hechos, que se active el indicador, la alarma, etc.

  Red 1: La temperatura y el indicador defectuoso puede ser un efecto común, por ejemplo si el indicador puede estar defectuoso, que a su vez activa la alarma que puede estar defectuosa como un efecto común.

  Red 2: La alarma y el indicador pueden ser causas comunes de la temperatura por que dependen de ella, por lo tanto ambos pueden tener sus defectos por medio de cadenas causales

  Red 3: La alarma puede ser un efecto común del indicador y de la temperatura, por lo que puede pasar que falle el indicador o la alarma por medio de una causa común

- De las redes dibujadas, escribe sus distribuciones conjuntas en términos de sus probabilidades condicionales.

  Las probabilidades conuntas son las siguientes, en orden de izquierda a derecha:

- $P(T, I, A, D_A, D_I) = P(T)P(D_I)P(I|D_I, T)P(A|D_I)P(D_A)$
- $P(T, I, A, D_A, D_I) = P(T)P(A|T)P(I|T)P(D_A|A)P(D_I|I)$
- $P(T, I, A, D_A, D_I) = P(T)P(I)P(A|T, I)P(D_A|A)P(D_I|A)$

- Asumiendo que las variables enteras I y T pueden tomar un máximo de 100 valores, ¿cuál sería el número de valores necesarios en cada nodo y el total en cada una de las redes?

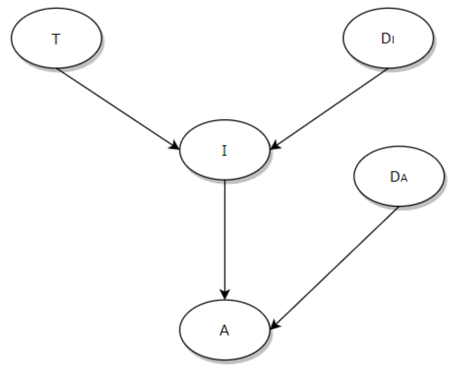


Figure 4: Red bayesiana 1

Dado que que el nodo I tiene dos padres, ocupamos todas las ombinaciones de los padres por las combinaciones de I.

$T = 100$
$D_I = 2$
$I = 100 * 100 * 2 = 20,000$
$D_A = 2$
$A = 2 * 2 = 4 \therefore$
$\#devalores = 100 + 2 + 20,000 + 2 + 2 = 20108$

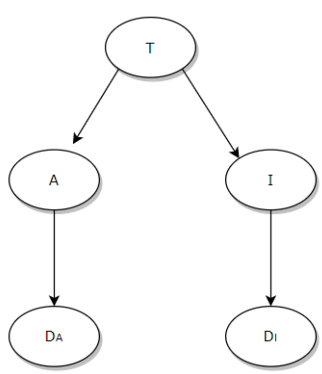


Figure 5: Red bayesiana 1

Como T no tiene padres, solo puede tomar 100 valores, pero I depende de T, entonces ocupamos las combinaciones, y de los demas nodos solo el numero de sus posibles valores

$T = 100$
$A = 2 * 100 = 200$
$I = 100 * 100 = 10000$
$D_A = 2 * 2 = 4$
$D_I = 2 * 100 = 200$
$\therefore$ el $\#devalores = 10504$

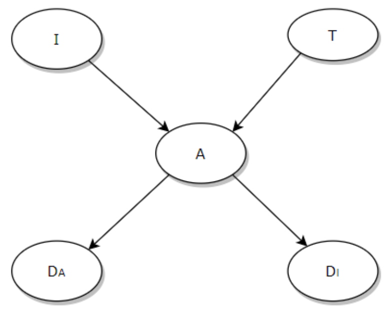


Figure 6: Red bayesiana 1

En esta red I y T no tienen padres, por lo tanto solo pueden tomar 100 valores, y las combinaciones de A serian el producto de ambas por su numero de valores.

$T = 100$
$I = 100$
$A = 2 * 100 * 100 = 20,000$
$D_A = 2 * 2 = 4$
$D_I = 2 * 2 = 4$
$\therefore$ el $\#devalores = 20208$

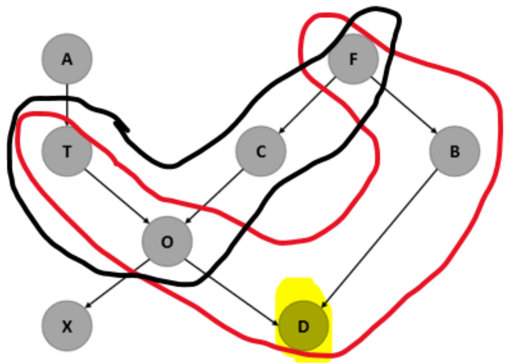
2

A una clínica le concierne el diagnóstico de enfermedades de pulmón. Como se puede ver en el modelo de la Figura 1, una visita a Asia (A) hace que la probabilidad de tener tuberculosis (T) aumente. Los nodos en la gráfica tienen el siguiente significado:

Etiqueta	**Significado**
A	Visita a Asia
T	Tuberculosis
C	Cáncer de pulmón
O	Tuberculosis o Cáncer de pulmón
F	Fumador
X	Prueba de rayos X positiva
B	Bronquitis
D	Disnea

Di si las siguientes relaciones de independencia condicional son verdaderas o falsas y explica por qué

- $T \perp\!\!\!\perp F|D$

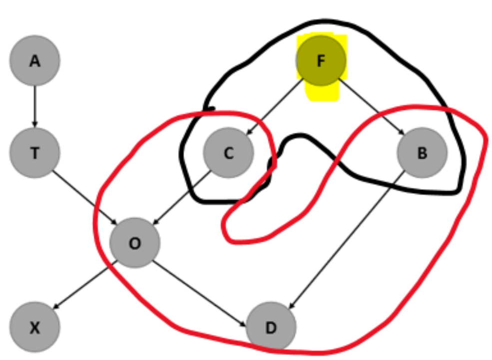


Caminos inactivos

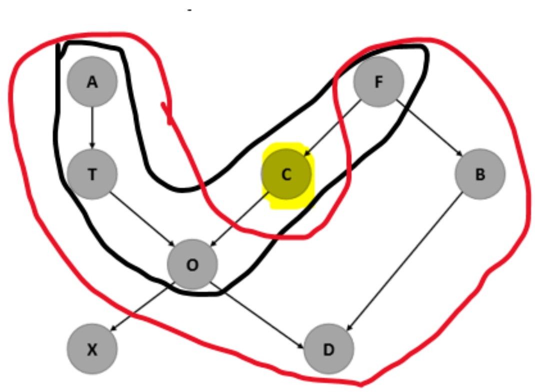
Caminos activos
$\{T,$ **O,C,F**$\}$
$\{$**T,O,D,B,F**$\}$

La tripleta que hace que **no garantize independencia** es un efecto comun $\{T, O, C\}$ ya que es un camino activo (tambien la tripleta {O,D,B} por efecto comun). Si encuentro un camino activo primero, omitiré poner los inactivos y poner los demas activos si existen mas de uno. Y en negritas denotare las tripletas que hacen que el camino sea activo o inactivo.

- $C \perp\!\!\!\perp B|F$



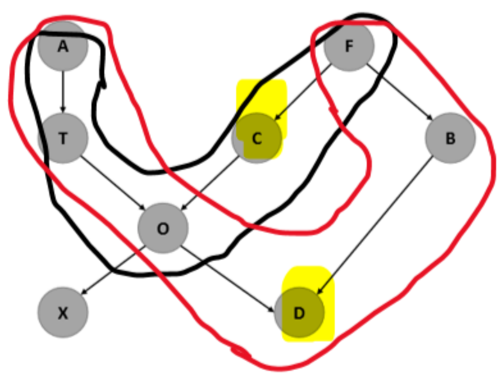
Caminos inactivos
$\{$**C,F,B**$\}$
$\{C,$ **O,D,B**$\}$
Se garantiza independencia condicional ya que ambos caminos son inactivos,



- $A \perp\!\!\!\perp F|C$

  Caminos inactivos
  $\{A, T,$ **O(X,D),C,F**$\}$
  $\{A, T,$ **O(X),D,B**$, F\}$
  Se garantiza independencia condicional.

- $A \perp\!\!\!\perp F|C, D$



  Caminos inactivos
  $\{A, T,$ **O(X,D),C,F**$\}$
  Camino activo
  $\{A, T,$ **O(X,D),C,F**$\}$
  No garantiza independencia.

3

Imagina una clínica que ayuda a pacientes con ébola en un área afectada por el virus. La red de la Figura 4 intenta capturar la dinámica por la cual las personas que sufren los síntomas pueden llegar a esta clínica y ver a un especialista. Existe la posibilidad de que alguien con ébola ($E = verdadero$) muestre síntomas, por ejemplo sangrado

($S = verdadero$) fiebre ($F = verdadero$) y visite la clínica ($V = verdadero$). El sangrado incrementa el riesgo de complicaciones ($C = verdadero$) y la persona puede ser llevada a ver un doctor especialista ($D = verdadero$).

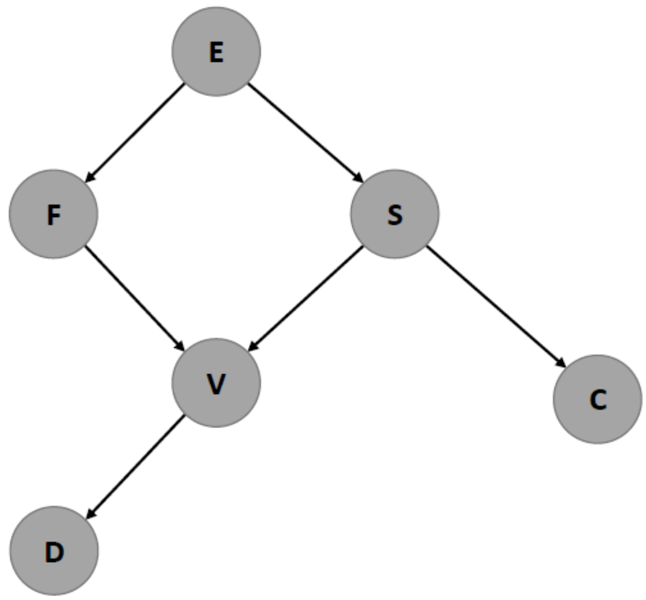


Figure 7: Modelo gráfico de detección de ébola

- $P(E = verdadero) = 0.01$
- $P(F = verdadero|E = verdadero) = 0.6$
- $P(F = verdadero|E = falso) = 0.1$
- $P(S = verdadero|E = verdadero) = 0.8$
- $P(S = verdadero|E = falso) = 0.05$
- $P(V = verdadero|F = verdadero; S = verdadero) = 0.8$
- $P(V = verdadero|F = verdadero; S = falso) = 0.5$
- $P(V = verdadero|F = falso; S = verdadero) = 0.7$
- $P(V = verdadero|F = falso; S = falso) = 0.0$
- $P(C = verdadero|S = verdadero) = 0.75$
- $P(C = verdadero|S = falso) = 0.1$
- $P(D = verdadero|V = verdadero) = 0.6$
- $P(D = verdadero|V = falso) = 0.0$

1. Escribe la distribución conjunta de la red bayesiana en función de las probabilidades condicionales.
   $P(E, F, S, V, C, D) = P(E)P(F|E)P(S|E)P(V|F, S)P(C|S)P(D|V)$

2. Si un paciente es llevado al doctor (D = verdadero), usando un paquete de software calcula la probabilidad de que no tenga ébola ($P(E = falso|D = verdadero)$).
   Jupiter Notebook
3. Convierte a la red bayesiana en un modelo gráfico no dirigido (campo aleatorio de Markov) y dibújalo. Captura tantas relaciones de independencia condicional como sea posible

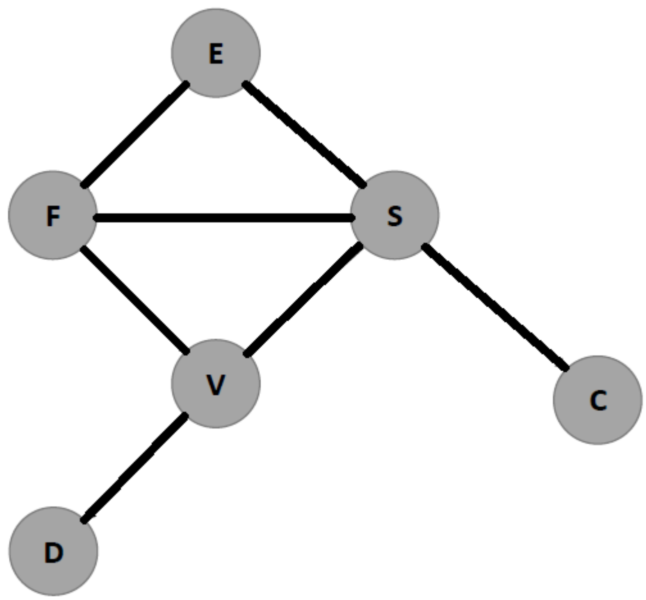


Figure 8: Campo aleatorio de markov

Solo moralicé los nodos F y S, siguiendo asi lo que entendí en clase que se tiene que asegurar que todos los nodos que aparecen en alguna probabilidad condicional de la distribucion, tienen que estar en por lo menos un click. Y uniendo F con S, todas las probabilidades condicionales que estan en la distribución aparecen en el click. (Me gustaría tener feedback para saber si es necesario que sea un click maximo). Por como entendi, no es necesario modificar las causas de efecto común. Fijandome en la dependencia condiconal entre $V \perp\!\!\!\perp C|S$

4. Debido a una campaña de concientización de la salud, las personas son alentadas a visitar la clínica en caso de que tengan fiebre. Esto incrementa la cantidad de visitas de personas confiebre sin importar el estado de cualquier otra variable.
   - ¿Qué probabilidades condicionales en la red se modifican debido a este cambio y en qué sentido?

   Si la distribucion conunta esta dada por:
   $P(E, F, S, V, C, D) = P(E)P(F|E)P(S|E)P(V|F, S)P(C|S)P(D|V)$ al momento de que las personas visitan la clinica en caso de fiebre, dado que conocemos V, las probabilidades que se modifican son $P(V|F, S)$ y por lo tanto pudiera ser que este efecto común activo, podemos decir que provoca que disminuya $P(V = Falso|F = Verdadero)$

- Describe cualquier efecto que esto tenga en la proporción de personas con complicaciones que visiten la clínica. Menciona exactamente qué probabilidades condicionales usaste para llegar a tu conclusión.

El camino C, S, V es una tripleta de causa común, dado que no conocemos informacion del nodo central S, podemos decir que es un camino activo por que tenemos informacion de ambos nodos hijos, Asi que estas probabilidades condicionales se pueden usar para determinar si una persona visita la clinica incluso aunque no presente sangrado.

Yo pienso que la proporcion de personas aumenta por lo mismo mencionado anteriormente.

5. Asume que alguien que no tiene fiebre va al doctor, ¿qué relación de independencia condicional existe en la distribución que no puede ser descubierta a través del grafo solamente?

La manera mas sencilla que se me ocurrió para atacar este problema es que pudiera justificar que la relacion de independencia condicional no puede ser descubierta dado como esta el grafo, para esto pudiera normalizar los nodos D, F, para que la cadena causal entre la tripleta $\{F, V, D\}$ no dependa solamente de V, en cuanto a la pura estructura del grafo. Intuitivamente pudiera ser que existiera otro síntoma relacionado con en ébola que no esta especificado en el grafo, dado que no podemos asegurar la independencia $F \perp\!\!\!\perp D|V$ si es que no tenemos informacion del nodo V.