

Introducción de redes bayesianas en el programa GeNIe

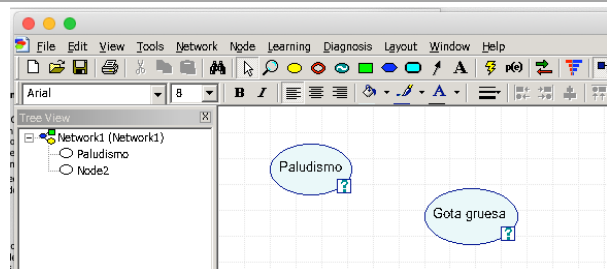
TUTORIAL

El programa GeNIe (Graphical Network Interface) nos permite crear redes bayesianas y redes de decisión de un modo intuitivo, utilizando un interfaz gráfico tipo *click-and-drop*. GeNIe es en realidad el interfaz gráfico de SMILE, un motor de inferencias bayesiano desarrollado en el laboratorio de Sistemas de Decisión de la Universidad de Pittsburgh. Puedes encontrar más información sobre Genie y SMILE en <https://www.bayesfusion.com>. Puedes descargar GeNIe desde el Campus Virtual, o si prefieres descargar la última versión, lo puedes hacer desde [esta web](#) (la versión académica es gratuita), pero ten en cuenta que algunos elementos del interfaz pueden ser diferentes. A continuación, vamos a explicar como definir un modelo basado en redes bayesianas con GeNIe y propondremos algunos ejercicios para que puedas practicar.

Definir una red con GENIE es muy sencillo, puesto que se hace utilizando un interfaz gráfico.

Para **definir un nodo**, simplemente se selecciona el botón con un óvalo y se coloca el nodo en el Panel de edición.

Por defecto, los nodos se crean como binarios (dos estados), y la probabilidad se distribuye uniformemente entre ellos.

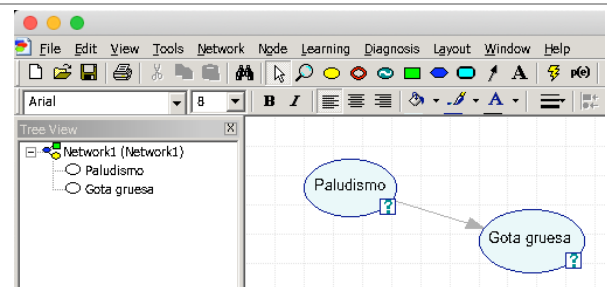


Para modificar las **propiedades de un nodo** podemos hacerlo de dos formas: o bien en el menú principal (*Node* → *Node properties*) o bien seleccionando el nodo y pulsando en el botón derecho del ratón). En la pestaña “General” podemos establecer el nombre, el identificador, el tipo de nodo, sus estados (es el momento de añadir más estados si es necesario, utilizando la opción “Add state”), los nombres de los estados, etc.

State name	State ID	Target	Default	Special name
ausente				
leve				
grave				

Una vez definidos los nodos y sus estados, el siguiente paso es reflejar las **relaciones de influencia causal** entre las variables (enlaces).

Para ello, pinchamos en el botón con una flecha, y en el panel de edición se inserta la flecha en la dirección adecuada



El siguiente paso es introducir los parámetros del modelo, es decir, las **probabilidades** a priori de los nodos sin padre y las probabilidades condicionadas de cada nodo dados sus padres.

Para ello, en la pestaña “definition” del nodo encontramos las tablas de probabilidad, que se construyen automáticamente teniendo en cuenta el número de estados del nodo, sus padres y el número de estados posibles de los padres. De esta forma, sólo hay que ir rellenando dicha tabla con las probabilidades condicionadas.

	no_paludismo	si_paludismo
no_paludismo	0.997	
si_paludismo	0.003	

	no_paludismo	si_paludismo
negativo	0.008	0.9994
positivo	0.992	0.0006

Estas probabilidades deben sumar 1 (por columnas). Cuando no sea así, aparecerá una señal de peligro amarilla. Para normalizar las probabilidades o bien para complementarlas a 1, podemos usar el botón de normalización, $\Sigma=1$ o el botón de complementar a 1, $1-\Sigma$.

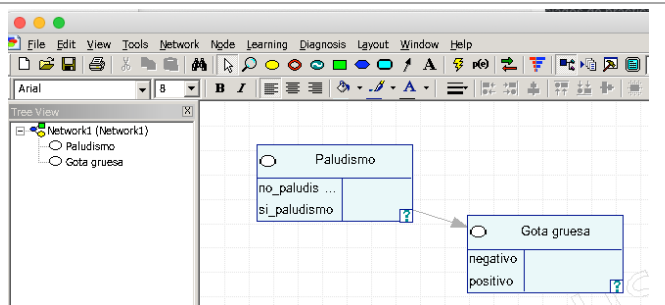
De este modo, si en la primera columna introducimos 8 y 992, seleccionamos la columna y presionamos el botón de normalizar, cambiará dichos valores por 0.008 y 0.992.

De manera similar, si en la segunda columna introducimos el valor 0.994, seleccionamos la segunda casilla y le damos a complementar, automáticamente cambiará el valor a 0.006. Podemos también seleccionar una fila entera para complementar a 1, y la calculará mediante $1 -$ (la suma de todos los valores en las casillas de la misma columna).

Finalmente, podemos *visualizar la red en modo "Bar chart"*, que nos permitirá ver no sólo sus estados, sino también los nodos y la evolución de las probabilidades. Para ello, seleccionamos la red completa y en el menú contextual veremos una opción "View as Bar Chart"

Paludismo	no_paludismo	si_paludismo
negativo	8	0.9994
positivo	9992	0.0006

Paludismo	no_paludismo	si_paludismo
negativo	0.0008	0.9994
positivo	0.9992	0



TAREA Y ENTREGA

Tarea: Para los siguientes enunciados, modela el problema como una red bayesiana e introdúcelo en GeNIe. Para los parámetros, en el caso de los nodos sin padres y, en ausencia de información, supón una distribución equiprobable para sus probabilidades a priori. En el caso de los nodos con padres, estima unos valores para las probabilidades condicionadas que sea acorde con el sentido común.

Entrega: Un único documento pdf, que para cada uno de los problemas contenga:

- Número de problema (no incluir el enunciado)
- Listado de nodos, con sus posibles valores
- Breve explicación del modelo propuesto
- Capturas de pantalla de los modelos y de las tablas de probabilidad (de todos los nodos con padres)

ENUNCIADOS

1. Una mañana, Ana decide visitar a su amiga Elena en su nuevo apartamento. Al entrar, Ana comienza a sentir picazón en los ojos. Inicialmente, piensa que dicha picazón podría deberse al síndrome del ojo seco. Sin embargo, al observar el apartamento, nota que hay varias plantas de interior. Ana considera entonces la posibilidad de que sus síntomas puedan deberse a una alergia al polen de las plantas.
2. Considera la siguiente situación: Sofía decide visitar al dermatólogo porque ha desarrollado una irritación en la piel de la cara. Durante la consulta, se produce la siguiente conversación:
 - *Sofía: Doctor, estoy preocupada por esta irritación que me ha aparecido en la piel de la cara. ¿Podría examinarme y decirme qué puede ser?*
 - *Dermatólogo: Por supuesto, Sofía. Cuénteme, ¿cuándo notó estos síntomas por primera vez?*
 - *Sofía: Apareció hace unos días. Al principio pensé que podría ser por el sol, ya que pasé mucho tiempo en la playa el fin de semana pasado.*
 - *Dermatólogo: Sí, la verdad es que puede deberse al sol. ¿Ha cambiado recientemente algún producto para el cuidado de la piel?*
 - *Sofía: Comencé a usar una nueva crema facial con retinol hace una semana. Leí que a veces puede causar irritación en la piel. ¿Cree que podría ser eso?*
 - *Dermatólogo: Es una posibilidad, aunque los síntomas que describe podrían indicar algo más. ¿Hay algo más que crea que deba saber?*
 - *Sofía: Bueno, me preocupa que pueda ser algo más serio. Mi tía tiene rosácea y he escuchado que puede ser hereditaria. ¿Cree que podría ser eso?*
 - *Dermatólogo: Es una preocupación válida, Sofía. La rosácea puede tener un componente genético y, efectivamente, puede manifestarse por primera vez en adultos. Los síntomas que describe son consistentes con esta condición.*
3. En el país Floravia se pueden encontrar varias especies de plantas con flor, llamémoslas *Floralis*, *Petalis* y *Ramosis*. Todas tienen un tamaño mediano, y sus hojas son o bien rugosas o lisas. Además, una observación atenta ha permitido deducir lo siguiente:
 - Todas las *Floralis* tienen flores blancas o rosas. Sus hojas son lisas, y de cada rama salen 4 o 5 ramas nuevas.
 - El número de ramas nuevas que salen de cada rama en las *Petalis* varía uniformemente entre 4 y 6, ambos inclusive. Sus hojas son rugosas.
 - En cuanto a las *Ramosis*, de cada rama salen 5 o 6 ramas nuevas, y sus flores son casi siempre rojas, pero a veces (20% de los casos) blancas.
 - Las plantas que en las que salen 4 ramas de cada rama se consideran robustas. El resto se consideran frágiles, a menos que el diámetro de la rama principal sea grande, lo cual ocurre en el 50% de los casos para las de 5 ramas y en el 80% para las de seis.
 - Plantea el problema de la clasificación de las plantas de Floravia mediante una red bayesiana

EJERCICIOS ADICIONALES (haz todos los que puedas)

4. Juan y Luisa llegan un día a casa y observan que la puerta está abierta. Piensan que puede haber habido un robo, pero entonces se dan cuenta de que el coche de su hija está aparcado en la calle y puede entonces ser que, al entrar en la casa, ella se haya dejado la puerta abierta. Al entrar en la casa, observan que los muebles de la entrada están tirados por el suelo con los cajones abiertos.
5. Juan está en la parada del autobús de la línea 20, y el autobús se está retrasando. Juan piensa que puede que haya retenciones de tráfico, pero también puede ser que el autobús haya sufrido una avería o que hayan suspendido el servicio de la línea por las obras del metro. El servicio de una línea se suspende cuando hay obras que la afectan y hay otras líneas en servicio que se pueden utilizar.
6. La policía está intentando establecer un modelo que permita razonar sobre los accidentes de tráfico causados por una pérdida de control del vehículo del conductor. Esta pérdida de control suele venir provocada por un error humano, una carretera resbaladiza, un fallo mecánico o un exceso de velocidad. El error humano suele deberse a una distracción del conductor y una capacidad de reacción mermada por alguna circunstancia (consumo de sustancias o cansancio). La carretera puede estar resbaladiza por vertido de sustancias o por las condiciones atmosféricas.
7. El problema de *Monty Hall*. A un concursante del concurso televisivo *Let's Make a Deal* se le pide que elija una puerta entre tres (todas cerradas), y su premio consiste en llevarse lo que se encuentra detrás de la puerta elegida. Se sabe que una de ellas oculta un coche, y las otras dos tienen una cabra. Una vez que el concursante ha elegido una puerta y le comunica al público y al presentador su elección, el presentador (que conoce en que puerta está el premio) abre una de las otras puertas y muestra una cabra. En este momento se le da la opción al concursante de quedarse con la puerta que eligió inicialmente o bien cambiar de puerta. ¿Debe el concursante cambiar de puerta, o mantener su elección?