

## Práctica de Ingeniería del Conocimiento

### El programa HUGIN

El programa HUGIN es la herramienta comercial más usada para el desarrollo y uso de redes bayesianas. La versión de demostración está disponible en <http://hugin.dk> (enlace Try\_it), donde también podemos encontrar mucha documentación relativa tanto al programa como a las redes bayesianas.

Definir una red con el programa HUGIN 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. Una vez definido el nodo, pueden añadirse más estados pinchando en el botón + o eliminarlos pinchando en -. Para definir los enlaces, pinchamos en el botón con una flecha, y en el panel de edición se inserta la fecha en la dirección adecuada.

Definidos los nodos, las tablas de probabilidad se construyen automáticamente teniendo en cuenta el número de estados del nodo, sus padres y el número de estados posibles de los padres, de forma que sólo hay que ir rellenando dicha tabla con las probabilidades condicionadas. Si los números introducidos no suman 1 (por columnas), HUGIN normalizará los resultados, de forma que si se quiere en una tabla de dos filas se puede introducir en una columna 6 y 4 con los mismos resultados que si se introduce 0.6 y 0.4.

Una vez definida la red, la compilamos (calculamos el estado inicial  $S_0$ ). Para ello, presionamos el botón que tiene un rayo y pasamos a un nuevo panel donde a la izquierda aparece una lista con los nodos y a la derecha la representación gráfica de la red.

Las probabilidades asociadas a cada nodo pueden visualizarse de dos formas distintas:

En la lista de los nodos. Para ello, presionamos el botón “Expand node list” que aparece justo encima de la lista, o bien nos situamos encima del cualquier nodo y pulsando el botón derecho seleccionamos “Expand list”. Estas dos opciones expanden la lista completa. Si sólo queremos expandir un nodo en particular, pulsamos dos veces con el botón izquierdo en el nodo de interés.

En el gráfico. Seleccionamos el conjunto de nodos cuyas probabilidades queremos ver, pulsamos el botón derecho del ratón y seleccionamos “Show monitor”. Aparece un monitor en el lugar del nodo que muestra las probabilidades.

En cualquier caso aparecen las probabilidades (multiplicadas por 100), junto con un diagrama de barras que ayuda a ver cómo se reparte la probabilidad entre los distintos estados posibles.

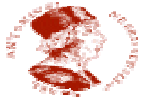
Para introducir evidencia, podemos hacerlo bien en la lista expandida o en el monitor de un nodo. Pulsamos dos veces con el botón izquierdo del ratón sobre el valor al que queremos instanciar la variable. Aparecerá una barra roja y la probabilidad de ese valor se pondrá a 100, y las del resto a 0. Una vez hecho esto, propagamos la evidencia. Para ello:

- Si lo que se desea es obtener la probabilidad de cada nodo dada la evidencia disponible, pulsamos el botón con el sumatorio.
- Si lo que se desea es obtener la combinación más probable, pulsamos el botón max.

También es posible realizar las siguientes operaciones:

- Volver al estado inicial de la red. Para ello se pincha el botón con una flecha redondeada.
- Eliminar una evidencia ya introducida. Para ello, seleccionamos “Retract evidence” del menú que aparece al pulsar con el botón derecho en la evidencia que queremos eliminar, y volvemos a propagar como se ha descrito anteriormente.

Para volver al modo de edición (para ver o modificar la red) se pincha en el botón en el que aparece un lápiz.



1. En el planeta Zyx se pueden encontrar varias clases de animales, llamemos a estas clases Wurros, Hobexas y Wackas. Todos tienen un tamaño muy pequeño, y sus pieles son o bien escamosas o bien están cubiertas de suave pelo. Además, una observación atenta ha permitido deducir lo siguiente:

- Todos los Wurros tienen 5 ó 6 patas. Su color es rojizo, y tienen la piel peluda y suave.
- El número de patas de las Hobexas es un entero que varía uniformemente entre 4 y 6, ambos inclusive. Su piel es escamosa.
- En cuanto a las Wackas, tienen 4 ó 5 patas, y ofrecen a la vista una tonalidad casi siempre azulada, pero a veces (20% de los casos) rojiza.
- Los animales que tienen un número impar de patas cojean siempre. Los animales que tienen un número par de patas cojean sólo cuando tienen alguna anomalía (malformación congénita, heridas, etc.), lo cual ocurre en el 10% de los casos para los animales de 4 patas, y en el 20% para los de seis.

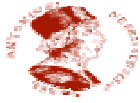
Se pide,

- a) Plantear el problema de la clasificación de animales de Zyx mediante un diagrama de influencia.
- b) Vemos un bicho rojizo que cojea. ¿Cómo lo clasificaremos?
- c) Las Hobexas y Wackas son confiadas e inofensivas. La escamosa piel de las Hobexas es muy apreciada, por lo que cada piel se vende por 6000 euros. La piel de las Wackas se vende por 4000 euros. Los Wurros no solamente son imposibles de capturar, sino que se defienden a codos, causando daños por valor de 2000 euros.

¿Vale la pena intentar capturar al animal avistado?

2. Una empresa de software se dedica a la planificación de proyectos y construcción de aplicaciones a diversos clientes. Con el fin de formular un modelo de la calidad de su proceso de desarrollo se consideran las siguientes variables junto con su dominio de valores posibles:

- A: capacidad de los jefes del proyecto <normal, elevada>
- B: capacidad de los programadores <normal, elevada>
- C: tecnología empleada en el desarrollo <normal, elevada>
- D: cumplimiento de la planificación <normal, elevado>
- E: número de errores <normales, elevados>
- F: nivel de ingresos <normales, elevados>
- G: calidad del software <normal, elevada>
- H: tiempo de desarrollo <bajo, medio, alto>
- I: satisfacción del cliente <bajo, medio, alto>



A	->	D
Normal		0.7 normal, 0.3 elevada
Elevada		0.6 elevada

B, C	->	E
Normal, normal		0.3 elevada
Normal, elevada		0.4 elevada
Elevada, normal		0.6 normal,
Elevada, elevada		0.4 normal

D,E	->	G
Normal, normal		0.5 elevada
Normal, elevada		0.7 elevada
Elevada, normal		0.2 normal,
Elevada, elevada		0.4 normal

D	->	F
Normal		0.6 normal,
Elevada		0.5 elevada

E	->	H
Normal		0.6 bajo, 0.3 medio
Elevada		0.1 bajo, 0.4 medio

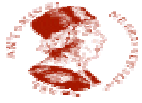
G	->	I
Normal		0.6 bajo, 0.3 medio
Elevada		0.1 bajo, 0.4 medio

Así por ejemplo, la primera línea de la primera tabla indica que:

$$p(D=\text{normal}|A=\text{normal})=0.7$$

$$p(D=\text{elevada}|A=\text{normal})=0.3$$

- ¿En qué medida influye la capacidad de los jefes de proyecto? Es decir, ¿a qué variables afecta la capacidad de los jefes y en qué medida? ¿Qué es más importante para el cliente: que mi empresa tenga buenos programadores o que tenga buenos jefes?
- ¿Y en el nivel de ingresos?
- ¿Cómo afecta la tecnología empleada?
- ¿Y el tiempo de desarrollo? Si no afecta, qué crees que deberíamos agregar a la red bayesiana para dar cuenta de este aspecto. Introduce los datos necesarios (estimados por ti) hasta que consigas una buena representación de esta influencia.
- Explica por qué y cómo se producen las separaciones direccionales. Para ello bloquea un camino (es decir, fija una evidencia) y di que variables se ven afectadas por este bloqueo. Es decir, al fijar una evidencia y modificar un nodo padre, ¿cómo varía la propagación de esa evidencia en todas las variables que eran antes dependientes?



## **INTRUCCIONES**

1. Se puede hacer de forma individual o en grupos 2 personas como máximo (en el caso de componerse el grupo por más de 2 personas se repartirá la nota a partes iguales)
2. La práctica debe consistir en una carpeta para cada ejercicio con las redes correspondientes y las probabilidades estimadas además de un documento con las respuestas y aclaraciones necesarias. Todo ello debe ser incluido en una carpeta con los nombres de los alumnos que han realizado la práctica (comprimida en un archivo zip) Deberá entregarse antes de las 23 horas del día 23 de Abril (plazo improrrogable: cualquier práctica recibida después no será evaluada) y únicamente se pueden entregar por correo electrónico a la dirección [cmalagon@nebrija.es](mailto:cmalagon@nebrija.es). Se enviará un acuse de recibo a la recepción del trabajo (de todas formas se recomienda guardar una copia de seguridad en casa)
3. Para la realización de la práctica se puede usar el programa Hugin o el simulador de redes bayesianas y diagramas de influencia del proyecto Elvira.