

Práctica final

Desarrollo de un Sistema Experto

Ingeniería del Conocimiento - GII
Curso 19-20



Universidad de Granada
Cumbreras Torrente, Paula
49087324-B

Índice de contenidos

● Resumen de cómo funciona el sistema experto	02
● Descripción del proceso seguido	03
○ Desarrollo de la base de conocimiento	
○ Validación y verificación del sistema	
● Descripción del sistema	05
○ Variables de entrada del problema, cómo se representan	
○ Variables de salida del problema, cómo se representan	
○ Conocimiento global del sistema	
○ Especificación de los módulos se han desarrollado	
■ Estructura en módulos	
■ Descripción de cada módulo	
○ Hechos y reglas de cada módulo	
● Breve manual de uso del sistema	11

Resumen de cómo funciona el sistema experto

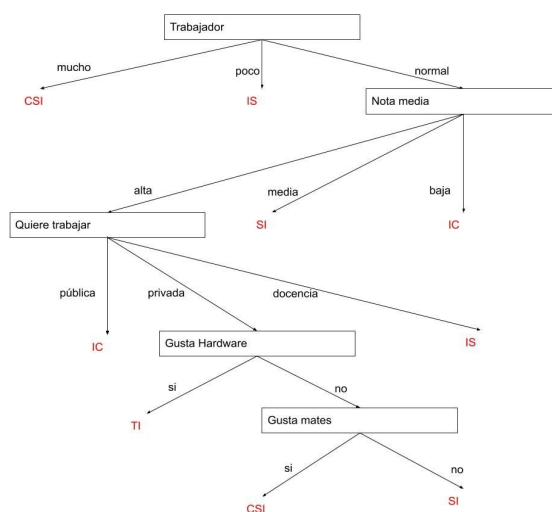
El sistema experto recomienda una serie de asignaturas del grado de Ingeniería Informática a un usuario en función de sus preferencias y cualidades. También cuenta con un módulo que recomienda una rama de la carrera. Para la recomendación de rama el sistema solo solicita responder a unas sencillas preguntas, a partir de las cuales razona y selecciona la rama más recomendada según los datos introducidos. Para la recomendación de asignaturas se necesita además de responder algunas preguntas, presentar una base de datos con las asignaturas posibles. Mediante un razonamiento probabilístico, el sistema seleccionará las asignaturas más recomendadas de entre todas las posibles. Tanto las ramas como las asignaturas incluirán un motivo en el que se justifique por qué han sido recomendadas. Ambos subsistemas funcionan como uno mismo, conectados por un módulo de control, que preguntará al principio qué recomendación busca el usuario; en función de la respuesta se lanzará el módulo de ramas o el de asignaturas.

Descripción del proceso seguido

Procedimiento seguido para el desarrollo de la base de conocimiento

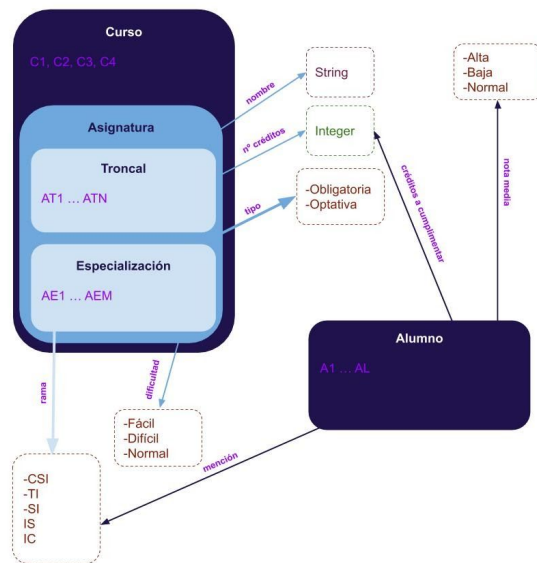
La tarea del sistema será recomendar a un usuario sobre qué asignaturas de la carrera escoger. El grado de Ingeniería Informática consta de 4 cursos: los 2 primeros tienen 10 asignaturas, el tercero tiene 30 y el cuarto tiene 43, por lo que el coste de la obtención de la información no supondría un impedimento aun cuando se trate de recomendar sobre todas las asignaturas de todos los cursos. Lo mismo para el coste de la salida: como mucho se podrán aconsejar 93 asignaturas, algo que no resulta demasiado complejo para un sistema de información. En un principio el sistema resulta viable.

Para la adquisición del conocimiento se ha realizado una entrevista abierta de tan solo una sesión incluida en formato pdf. La viabilidad del sistema propuesto queda expresada en dicho documento. Además se plantea el desarrollo de un subsistema que asesore al usuario en la selección de la rama de la carrera en la que especializarse. Para dicho subsistema se ha generado un árbol de clasificación a partir de las recomendaciones para la selección de rama por parte del experto:



Las recomendaciones del experto, la generación del árbol y las reglas de decisión obtenidas han sido añadidas en un fichero “Obtencion_conocimiento” de tipo pdf.

Una vez habiendo obtenido el conocimiento del experto podemos empezar a estructurar el conocimiento, teniendo en cuenta las variables de entrada, los datos que manejará el sistema y las variables de salida. Sabemos entonces que contaremos con un cierto número de asignaturas de cada curso, que pueden ser troncales o de especificación, con un nombre, un número de créditos, dificultad y obligatoriedad. Para formar el consejo pediremos al usuario que nos dé información sobre el número de créditos que desea cursar, su nota media actual y su rama de especialización. Esto ha sido representado en la siguiente ontología:



Ya tenemos el modelo conceptual, para establecer el modelo formal definimos lo que serán nuestros módulos: sistema de recomendación de ramas y sistema de recomendación de asignaturas, ambos con estructura de sistema simple de análisis en paralelo (parten de una serie de opciones, las analizan paralelamente y elaboran una respuesta). Aunque los subsistemas funcionen independientemente, la respuesta del primero puede influenciar en las decisiones que tome el usuario a la hora de introducir los datos solicitados en el segundo subsistema, por lo que se recomienda ejecutarlos en dicho orden.

Procedimiento de validación y verificación del sistema

Verificación del sistema:

- Consistencia: El sistema no llega a conclusiones incoherentes. Para lograr ésto se han incluido reglas de control que tratan las inconsistencias (evitando redundancia, contradicciones...) y controlan el orden de ejecución. Para evitar incoherencias de reglas, por ejemplo, siempre que se quiera modificar un valor de un hecho, primero se eliminará por completo y luego se volverá a crear ya modificado.
- Precisión: Es sintácticamente correcto. Ambos subsistemas pueden ser ejecutados en el IDE correspondiente sin que presenten ningún error al cargarse.
- Completitud: El sistema sabe actuar en todos los casos. Los únicos momentos en los que no realice una recomendación serán momentos en los que la información sea realmente insuficiente; por ejemplo, en el caso del sistema de recomendación de asignaturas, porque no se haya seleccionado una base de datos de la que obtener la información de las asignaturas.

Validación del sistema:

- El sistema cumple las especificaciones del modelo de diseño.
- Comunicación entre subsistemas.
- Interfaz comprensible para el usuario.
- Tiempos de ejecución despreciables.

Descripción del sistema

Variables de entrada del problema, cómo se representan

Las variables de entrada para la recomendación de rama se recogen a través de una entrevista inicial que consta de 5 preguntas. Se obtendrá una variable por pregunta:

- Trabajador: variable que puede tomar los valores “mucho”, “poco” o “normal” dependiendo de lo trabajador que se considere el usuario.
- Nota media: variable que puede tomar los valores “alta”, “media” o “baja.”
- Preferencia de trabajo: variable que representa el sector en el que desea trabajar el usuario, puede tomar los valores “docencia”, “empresa_privada” o “empresa_publica.”
- Gusto hardware: variable que indica si al usuario le gusta el hardware o no.
- Gusto matemáticas: variable que indica si al usuario le gustan las matemáticas o no.

Todas estas variables son opcionales, no será necesario responder a todas las preguntas, aunque es recomendado realizarlo para obtener un consejo más acertado.

Para la recomendación de asignaturas se piden las siguientes variables:

- Base de datos de asignaturas: consiste en un archivo en formato txt donde se representan las asignaturas junto a algunas de sus características (nombre, rama, obligatoriedad, cuatrimestre, dificultad y créditos, en éste orden). Se han añadido 4 bases de datos con las asignaturas agrupadas por curso. También se podrá incluir una nueva base de datos, añadiendo manualmente el nombre del archivo.
- Rama: variable que almacena la rama escogida por el usuario.
- Nota media: variable que puede tomar los valores “alta”, “media” o “baja.”
- Cantidad de créditos y número de asignaturas por cuatrimestre: variables que almacenan el número total de créditos a cumplimentar (ha de ser múltiplo de 6) y cantidad de asignaturas del primer cuatrimestre (ha de ser menor que el número total de créditos entre 6).

Estas variables vuelven a ser opcionales, aunque, como es evidente, si no se introduce una base de datos no se podrá recomendar ninguna asignatura.

Variables de salida del problema, cómo se representan

Las variables de salida, tanto en la recomendación de rama como en la de asignaturas, consiste en un hecho Consejo, que almacena el nombre de la rama o asignatura a recomendar, el motivo para ello y el apodo del experto. Para la recomendación de ramas se recomendará 1 o 2 ramas como máximo; para la recomendación de asignaturas se recomendarán tantas asignaturas como se indique en las variables de entrada cantidad de créditos y número de asignaturas por cuatrimestre.

Conocimiento global del sistema

El sistema de recomendación de ramas ha de conocer los siguientes hechos y relaciones:

- Inicialización de las ramas con su nombre completo.

- Equivalencia entre el nombre de las ramas y su abreviatura.
- Inicialización de los consejos para cada rama, con el motivo vacío y “Clips” como nombre de experto. Esto último se ha usado para indicar que el consejo no ha sido realizado por el experto, sino simplemente inicializado por el sistema.
- Inicialización de la puntuación de cada rama a 0, usado para el razonamiento.
- Inicialización de las preguntas, indicando que no han sido respondidas.
- Equivalencia entre el número de la pregunta y la variable de entrada que se obtiene.
- Inicialización de justificaciones que consisten en un texto relacionado con las variables de entrada que se añadirá al consejo si resulta conveniente.
- Inicialización de los valores de las preguntas, es decir, cómo influye cada respuesta dada a cada una de las preguntas en la elección de cada una de las ramas; será usado para el razonamiento.
- Inicialización del proceso de obtención de datos para controlar el orden de ejecución.

El sistema de recomendación de asignaturas únicamente conocerá los siguientes hechos y relaciones:

- Inicialización de las probabilidades para el razonamiento; no todas se cargarán inicialmente, algunas se irán añadiendo según se vayan obteniendo las variables de entrada. En nuestro sistema solo se cargan inicialmente las probabilidades correspondientes a la obligatoriedad (se recomendará siempre antes las asignaturas obligatorias que las optativas).
- Inicialización de las preguntas y del proceso de obtención de datos.

Especificación de los módulos se han desarrollado

Estructura en módulos

Como ya se ha indicado anteriormente, el sistema consiste en dos subsistemas de recomendación almacenados en ficheros independientes: uno para recomendar rama y otro para recomendar asignaturas. Ambos se dividen en 3 secciones, correspondientes a la obtención de datos, el razonamiento y el fin del razonamiento. Se ha implementado un tercer módulo usado para la conexión de ambos subsistemas llamado control.

Descripción de cada módulo

Descripción del sistema de recomendación de ramas:

- Obtención de datos: Consiste en una entrevista de 5 preguntas. El usuario podrá elegir no responder a una pregunta o incluso dar por finalizada la entrevista antes de llegar a la última. Las respuestas se almacenarán en variables pues influirán en el proceso de razonamiento. Si el usuario responde “ns” a una pregunta, una vez se finalice la entrevista, ésta respuesta será eliminada al no tratarse de una respuesta útil.
- Razonamiento: Se ha empleado un sistema de razonamiento basado en puntos. Todas las ramas comenzarán con una puntuación de 0, que indica el grado de afinidad que tiene el usuario con cada una de ellas. De acuerdo con el árbol de clasificación mostrado anteriormente, cada respuesta otorgada aumentará la

puntuación de las ramas en cierto grado, hasta un máximo de 100. Si las respuestas han aumentado la puntuación de una rama, se incluirá una justificación en el motivo del consejo.

- Fin del razonamiento: Se recomendará aquella rama con la que el usuario tenga mayor grado de afinidad. En caso de empate se recomendarán 2.

Descripción del sistema de recomendación de asignaturas:

- Obtención de datos: Consiste en una entrevista de 4 preguntas. El usuario podrá elegir no responder, pero la primera pregunta, correspondiente a la base de datos, ha de ser respondida para poder obtener una recomendación. La probabilidad de que se recomiende una asignatura o no se ve afectada por las respuestas dadas; por ejemplo, si se elige la rama CSI, la probabilidad de que el sistema recomiende una asignatura de la rama IC será de 0.1.
- Razonamiento: Se lleva a cabo un razonamiento probabilístico.
- Fin del razonamiento: Se seleccionarán las n asignaturas del primer cuatrimestre y las m del segundo con más probabilidad de ser recomendadas, siendo n y m los números de asignaturas del primer y segundo cuatrimestre respectivamente. Se tendrán en cuenta las asignaturas de 12 créditos, que contarán como 2 asignaturas simples.

Descripción del sistema de control: obtiene la única función de ofertar los subsistemas anteriormente mencionados y controlar la ejecución de los mismos. Cada módulo podrá ejecutarse varias veces, para salir del sistema se ejecutará el módulo null.

Hechos y reglas de cada módulo

Sistema de recomendación de ramas:

- Hechos para representar los consejos y las ramas:
 - (Rama <nombre de la rama>) representa las posibles ramas a recomendar.
 - (Consejo_Rama <nombre de la rama> <texto del motivo> <apodo del experto>) representa que la rama <nombre de la rama> ha sido recomendada por el experto <apodo del experto> bajo el razonamiento <texto del motivo>.
 - (Es_Rama <rama> <nombre>) relaciona la rama con abreviatura <rama> con el <nombre> verdadero de dicha rama.
 - (Puntuacion <rama> <puntuacion>) almacena un valor <puntuacion> que irá aumentando en función al grado de afinidad del alumno con la rama <rama>.
- Hechos para representar las preguntas y respuestas:
 - (Pregunta_Rama <orden> <respuesta util> <puntuada>) es una representación general de todas las preguntas, donde <orden> indica a cuál de todas se refiere. Si la respuesta obtenida es util, se mostrará en

- <respuesta util> y si esta pregunta ya ha tenido repercusión sobre las puntuaciones de las ramas, <puntuada> lo reflejará.
- (Respuesta <tipo> <respuesta>) es un template general para todas las respuestas; en <tipo> se indica a qué responde y en <respuesta> se almacena lo respondido por el alumno.
 - (Es_Pregunta <tipo> <orden>) relaciona el <tipo> de respuesta dada con el número de la pregunta en la que se responde.
 - (Justificaciones <orden> <texto>) almacena el texto que incluirá el motivo del consejo si la pregunta <orden> ha influenciado en la recomendación.
 - (Valor_Pregunta <rama> <pregunta> <respuesta> <puntos>) almacena que la respuesta <respuesta> de la pregunta número <pregunta> aumenta en <puntos> la puntuación o grado de afinidad de la rama con abreviatura <rama>.
- Hechos para controlar el orden de ejecución:
 - (Obtencion <continuar>) continuar tomará el valor S hasta que finalice el proceso de obtención de datos del usuario.
 - (Razonamiento <finalizado>) finalizado tomará el valor N hasta que se finalice el razonamiento, es decir, hasta que se genere un consejo.
 - (Selección <seleccionado>) seleccionado tomará el valor S cuando se haya seleccionado la rama a aconsejar tras realizar el razonamiento.
 - Reglas para la obtención de datos:
 - (trabajador) regla para hacer la primera pregunta.
 - (terminar_primera) regla para dar por finalizada la primera pregunta.
 - (nota) regla para hacer la segunda pregunta.
 - (terminar_segunda) regla para dar por finalizada la segunda pregunta.
 - (futuro) regla para hacer la tercera pregunta.
 - (terminar_tercera) regla para dar por finalizada la tercera pregunta.
 - (hardware) regla para hacer la cuarta pregunta.
 - (terminar_cuarta) regla para dar por finalizada la cuarta pregunta.
 - (mates) regla para hacer la quinta pregunta.
 - (terminar_entrevista) regla para terminar la quinta pregunta y la entrevista.
 - (respuesta_inutil) regla para borrar información innecesaria.
 - Reglas para el razonamiento:
 - (calidad_respuesta_rama) regla para avisar de la posible poca precisión en la recomendación por falta de datos.
 - (puntuar) regla para estudiar el incremento que suponen las respuestas en la puntuación de cada rama.
 - (justificar) regla para añadir al una nueva justificación al motivo.
 - (consejo_valido) regla para marcar un consejo como válido.

- (fin_razonamiento) regla para asignar el grado de afinidad obtenido y finalizar el proceso de razonamiento.
- Reglas para el fin del razonamiento:
 - (mejor_rama) regla para seleccionar la mejor rama aconsejada en función al grado de afinidad.
 - (empate_ramas) regla para seleccionar las dos mejores ramas aconsejadas en caso de que haya un empate en el grado de afinidad.
 - (final) regla para mostrar por pantalla el consejo final.
 - (error final) regla para indicar que no se ha podido realizar el razonamiento por falta de información.

Sistema de recomendación de asignaturas:

- Hechos para representar los consejos y las asignaturas:
 - (Asignatura <nombre> <rama> <tipo> <cuatrimestre> <dificultad> <creditos> <probabilidad>) representa una asignatura y todas sus características. Todos los datos han de aparecer en el fichero de lectura, salvo <probabilidad> que será usada por el sistema para el razonamiento.
 - (Consejo <nombre de asignatura> <texto del motivo> <apodo del experto>) representa que la asignatura <nombre de asignatura> ha sido recomendada por el experto <apodo del experto> bajo el razonamiento <texto del motivo>.
 - (Probabilidad <tipo> <valor> <probabilidad>) almacena un valor <probabilidad> correspondiente a lo que el valor <valor> de la característica <tipo> aporta a la selección de asignaturas. Algunos de estos valores serán iniciales, otros variarán en función de las preferencias del usuario.
- Hechos para la lectura, las preguntas y el control:
 - (Pregunta <orden> <finalizada>) es una representación general de todas las preguntas, donde <orden> indica a cuál de todas se refiere. Si la pregunta ya ha sido respondida, se respresentará en <finalizada>.
 - (READ <archivo> <nombre>) almacena en <archivo> el nombre del fichero de lectura y en <nombre> se almacena un nombre lógico.
 - (Leyendo <archivo> <continuar>) continuar valdrá S si aún no se ha alcanzado el fin del archivo <archivo>.
 - (Obtencion <continuar>) continuar tomará el valor S hasta que finalice el proceso de obtención de datos del usuario.
 - (Creditos <primer cuatrimestre> <segundo cuatrimestre>) reparte el número de créditos entre los dos cuatrimestres, de forma que se pueda controlar el número de recomendaciones por cuatrimestre.
- Reglas para la obtención de datos:
 - (selección_base_datos) regla para realizar la primera pregunta y seleccionar una o varias bases de datos.

- (nueva_base) regla para añadir una nueva base de datos.
- (recoger_datos) regla para abrir un archivo.
- (nueva_asignatura) regla para leer el archivo y añadir una asignatura al sistema.
- (fin_archivo) regla para cerrar un archivo.
- (seleccion_rama) regla para realizar la segunda pregunta y ajustar probabilidades a la respuesta.
- (seleccion_nota) regla para realizar la tercera pregunta y ajustar probabilidades a la respuesta.
- (seleccion_creditos) regla para realizar la cuarta pregunta.
- Reglas para el razonamiento:
 - (calidad_respuesta) regla para avisar de la posible poca precisión en la recomendación por falta de datos.
 - (puntuar) regla para estudiar la probabilidad de cada asignatura aún no aconsejada.
- Reglas para el fin del razonamiento:
 - (mejores_primer_cuatrimestre) regla para mostrar por pantallas tantos consejos de asignaturas del primer cuatrimestre como se haya indicado en la cuarta pregunta.
 - (mejores_segundo_cuatrimestre) regla para mostrar por pantallas tantos consejos de asignaturas del segundo cuatrimestre como se haya indicado en la cuarta pregunta.
 - (error_base_datos) para indicar que no se ha podido realizar el razonamiento ya que no se ha especificado ninguna base de datos.

Sistema de control:

- (Modulo <nombre>) controla los módulos anteriores, pudiendo tomar <nombre> los valores ramas, asignaturas o null.
- (seleccion) regla para seleccionar el módulo a ejecutar.
- (terminar_modulo) regla para terminar la ejecución de un módulo.

Breve manual de uso del sistema

1. Se han añadido 4 bases de asignaturas al sistema; si se desea añadir una nueva, ésta ha de estar en formato txt y seguir la nomenclatura (Nombre Rama Obligatoriedad Cuatrimestre Dificultad Créditos) donde
 - a. Nombre ha de ser un símbolo sin espacios.
 - b. Rama ha de tomar uno de los siguientes valores: troncal CSI IS IC SI TI.
 - c. Obligatoriedad ha de tomar uno de los siguientes valores: O V.
 - d. Cuatrimestre ha de tomar uno de los siguientes valores: 1 2.
 - e. Dificultad ha de tomar uno de los siguientes valores: difícil fácil normal.
 - f. Créditos ha de tomar uno de los siguientes valores: 6 12.
2. Cargar los ficheros “control.clp,” “ramas.clp” y “asignaturas.clp” en dicho orden. Ejecutarlos.
3. Se pedirá seleccionar un módulo
 - a. Para el módulo ramas, se pedirá lo siguiente:
 - i. Responder a una pregunta con alguna de las opciones indicadas. Existirá una opción “ns” para introducir en caso de que no se conozca la respuesta o se decida no responder.
 - ii. Confirmar si se desea seguir realizando la entrevista. En caso negativo se finaliza la entrevista; de lo contrario, se repite el proceso hasta que se hayan respondido todas las preguntas restantes.
 - b. Se mostrará la(s) rama(s) recomendada(s).
 - c. Para el módulo asignaturas, se pedirá lo siguiente:
 - i. Seleccionar una de las bases de datos que ya habían sido creadas. Se podrá seleccionar más de una.
 1. 1 corresponde a “asignaturas_primer.txt”
 2. 2 corresponde a “asignaturas_segundo.txt”
 3. 3 corresponde a “asignaturas_tercero.txt”
 4. 4 corresponde a “asignaturas_cuarto.txt”
 - ii. Si se desea añadir otra base de datos, indicar el nombre del archivo.
 - iii. Indicar la rama escogida o recomendada por el sistema.
 - iv. Indicar la nota media.
 - v. Indicar el número de créditos a cumplimentar y el número de asignaturas del primer cuatrimestre. Esta pregunta se repetirá hasta que los datos sean válidos: el número de créditos ha de ser múltiplo de 6 y el número de asignaturas no puede superar el total de créditos entre 6.
 - d. Se mostrarán tantas asignaturas recomendadas por cuatrimestre como se haya indicado.
4. Para salir del sistema, introducir null.