

Inteligencia Artificial
Práctica Sistemas Basados en el Conocimiento -
LibrosIA

Albert Lobo, Ferran Pelegrina, Jaime Pascual

Diciembre 2015

Índice

1. Introducción	6
2. Identificación	7
2.1. Descripción del problema	7
2.2. Viabilidad de la solución	8
2.3. Fuentes del conocimiento	8
2.4. Objetivos	9
3. Conceptualización	9
3.1. Descripción de los conceptos del dominio	9
3.2. Descripción detallada de los problemas y subproblemas que in- tervienen en la resolución	11
3.3. Ejemplos del conocimiento experto extraído del dominio	13
3.4. Descripción informal del proceso de resolución, y cómo se organi- zan los problemas y subproblemas	14
4. Formalización	14
4.1. Se explica detalladamente como se ha construido la ontología . . .	14
4.2. La ontología incluye todos los conceptos, todos los atributos, y todas las relaciones necesarias	15

4.3.	Documentación detallada de la ontología (se describen todos los elementos de la ontología) incluyendo un grafo de la jerarquía de conceptos	16
4.4.	Se detalla el proceso de razonamiento de cada subproblema identificado	17
4.5.	Justificación de la metodología de resolución de problemas más adecuada y como los problemas identificados en la fase anterior encajan en la metodología	19
5.	Implementación	20
5.1.	Construcción de la ontología	20
5.2.	Construcción de la base de datos	20
5.3.	Construcción del programa	21
5.3.1.	Templates	21
5.3.2.	Variables	22
5.3.3.	Hechos	22
5.3.4.	Funciones	23
5.3.5.	Módulos	24
5.4.	Versiones	28
5.4.1.	Inicial	28
5.4.2.	Aumentado 1	28
5.4.3.	Aumentado 2	29

5.4.4. Definitivo	29
6. Pruebas	29
6.1. Prueba 1	29
6.1.1. Introduccion	29
6.1.2. Resultado Esperado	30
6.1.3. Entrada CLIPS	30
6.1.4. Sortida CLIPS	32
6.1.5. Resultado	32
6.2. Prueba 2	33
6.2.1. Introduccion	33
6.2.2. Resultado Esperado	33
6.2.3. Entrada CLIPS	33
6.2.4. Salida CLIPS	35
6.2.5. Resultado	35
6.3. Prueba 3	36
6.3.1. Introduccion	36
6.3.2. Resultado Esperado	36
6.3.3. Entrada CLIPS	36
6.3.4. Salida CLIPS	38

6.3.5. Resultado	38
6.4. Prueba 4	38
6.4.1. Introduccion	38
6.4.2. Resultat Esperado	39
6.4.3. Entrada CLIPS	39
6.4.4. Salida CLIPS	40
6.4.5. Resultado	41
6.5. Prueba 5	41
6.5.1. Introduccion	41
6.5.2. Resultado Esperado	41
6.5.3. Entrada CLIPS	42
6.5.4. Salida CLIPS	43
6.5.5. Resultado	44
7. Conclusión y mejoras	44

1. Introducción

Este documento es el informe sobre el desarrollo de la práctica de SBC (Sistemas Basados en el Conocimiento) del cuatrimestre de otoño 2015-2016. En ésta debemos implementar un sistema experto que recomienda libros (en particular, de novelas de ficción) a partir de las preferencias, gustos, y costumbres de lectura del usuario.

El sistema contiene una base de datos con libros ficticios -lo bastante grande para dar una amplia variedad de recomendaciones a los diferentes perfiles de usuario-, de la que seleccionará los 3 libros como recomendaciones.

Hemos utilizado la metodología en cascada para desarrollar este SBC. Esta consiste en 5 fases:

- Identificación del problema
- Conceptualización
- Formalización
- Implementación
- Prueba

A continuación describiremos con más detalle cómo hemos realizado cada una de ellas en este Sistema Experto de recomendación de libros.

2. Identificación

2.1. Descripción del problema

El problema consiste en la recomendación de libros a los lectores. Encontrar una solución a recomendar 3 libros que se adapten a las necesidades del usuario y que esté basada en información relevante en el mundo de los libros. Por tanto debemos encontrar una solución que se comporte como lo haría un experto en el tema de los libros y que a su vez conociera al usuario. Dado que este dominio es muy amplio hemos decidido centrarnos en el género y los subgéneros de ficción.

El volumen de información a procesar es muy alto, el amplio abanico de la literatura, los gustos de una persona, las características físicas a la hora de comprar un libro,... Esto ya nos da una idea de que clase de técnica hemos de utilizar a la hora de resolver nuestro problema.

Hemos distinguido los datos relevantes del usuario en 3 categorías fundamentales:

Preferencias explícitas: Se trata de los datos que el usuario nos da explícitamente sobre sus gustos; por ejemplo, su temática preferida (ciencia-ficción, fantástica, novela romántica o negra). Sin embargo, existen una gran variedad de parámetros que podríamos tener en cuenta, así que hemos optado por dar a cada temática una serie de atributos propios, no compartidos con libros del resto de temáticas.

Datos personales: La edad del lector, por ejemplo, juega un papel importante, ya que no queremos recomendar libros para adultos a menores. Al igual que la edad, existen más datos que podemos utilizar para dar buenas recomendaciones: tiempo de lectura al día, lugar donde el usuario lee,...

Conocimiento por similitud: También consideramos que algunas reco-

mendaciones no se pueden hacer más que por el propio conocimiento de los datos particulares del sistema. Por ejemplo, si tenemos una trilogía de libros, y una de las entregas es el libro favorito del lector, posiblemente los otros libros de la serie (o incluso una “versión cutre”) sean del interés del lector.

Con estos datos podemos determinar qué atributos de los libros debemos buscar en las recomendaciones que recibirá el usuario.

Para hacer el problema más asequible, utilizamos una base de datos ficticia, cuyo propósito era tener recomendaciones para una amplia variedad de perfiles de lector, sin cargar la BD con cientos de libros.

2.2. Viabilidad de la solución

Nos encontramos ante un problema de búsqueda: Necesitamos recomendar 3 libros de entre los que tenemos en el sistema. Naturalmente, sin conocimiento esta labor sería un ejercicio de futilidad, ya que las recomendaciones resultarían inadecuadas la mayor parte de veces.

Sin embargo, disponemos de información acerca de las características que definen un libro, y también sabemos cómo relacionarla con los gustos, costumbres, y datos del usuario -por lo que, mediante un SBC, podemos efectivamente encontrar libros que serán recomendaciones buenas para el lector.

2.3. Fuentes del conocimiento

El conocimiento lo tenemos en las librerías, libreros, críticos, internet y por su puesto en los libros. Cada libro tiene unas características únicas, una temática, una trama de una complejidad x , un tamaño,... Tenemos que extraer ese conocimiento y transmitirlo a nuestro sistema experto. Por otro lado tenemos que extraer conocimiento de los expertos en el dominio, un librero, un crítico o

uno o más expertos en el tema. De ellos debemos extraer información sobre la relación entre los libros. Qué características de qué libros son más relevantes y cuáles están más relacionadas con las de otros libros. Por ejemplo la temática de un libro, ¿si a una persona le gusta un libro, es probable que le gusten otros de temática similares? ¿Cómo de probable?

Para crear este sistema basado en el conocimiento necesitamos la información de librerías y bibliotecas. Los libros y sus características. Posteriormente necesitamos la sabiduría de uno o dos expertos en la materia que nos puedan decir cuales son las características más importantes (el autor, la temática, la trama,...)

2.4. Objetivos

El objetivo de este trabajo es recomendar a una persona un libro en base a sus gustos. en base a unas preguntas que se le hará aun usuario poder recomendarle un libro de su agrado. De manera que si el usuario no sabe qué leer debemos extraerle la información y a raíz de ella elaborar una lista de posibilidades, cuanto más información le extraemos más certera será nuestra solución.

3. Conceptualización

3.1. Descripción de los conceptos del dominio

El elemento de nuestro dominio es el libro, al que hemos extraído las características que nos han parecido más importantes para poder clasificar a nuestros usuarios.

Por tanto un libro consta de un autor, un título, un atributo best seller,

la edad recomendada y por último la longitud.

Características del **Libro**:

- El autor del libro
- El título del libro
- Si el libro fue un best seller o no
- La edad recomendada para leer el libro. Si es un libro para todos los públicos o solo para mayores de edad.
- La longitud del libro discretizada en corto o largo. Se puede ver como de bolsillo o no.
- Temática
- Si el libro pertenece a una obra clásica o no
- El formato del libro, de entre los cuales identificamos el formato ebook, el de tapa dura y el de tapa blanda.
- Tipo lector casual o avanzado.
- Libros similares, que es un listado de libros relacionados con este.

Otro de nuestros elementos del dominio son los libros específicos a un sub género. Contienen atributos extras pertenecientes a cada uno de estos sub géneros. En algunos casos son temáticas habituales dentro del subgénero, en otros, se trata de los personajes y/o elementos típicos dentro de una trama de su tipo. Los sub géneros con los que trabajamos son la fantasía, la novela negra, la ciencia ficción y la novela romántica.

Características del **Libro fantasía**:

- Razas en las que se centra la historia de libro. Tales como enanos, elfos, humanos,...

- Si la trama ocurre en un mundo completamente inventado o si está basada en la tierra.

Características del **Libro novela negra**:

- Si hay acción en el libro o no.
- Que tipo de personaje es el protagonista. Un detective, un mafioso, o alguien perteneciente a los juzgados (juez o abogado) protagonistas

Características del **Libro ciencia ficción**:

- El libro trata de una utopía o una distopía.
- En el libro hay futuros alternativos
- El libro trata de viajes en el tiempo

Características del **Libro novela romántica**:

- El tipo de romance. Es un romance histórico, contemporáneo o fantástico.
- Es muy sensual o no

3.2. Descripción detallada de los problemas y subproblemas que intervienen en la resolución

Para poder lograr una correcta solución debemos identificar previamente los subproblemas que engloba nuestro problema principal, y ver cómo podemos solucionarlos mediante un SBC. Los subproblemas que identificamos son la recopilación de información, lo cual conlleva una serie de subproblemas, valorar la información en base al conocimiento que recopilamos y por último imprimir la solución o soluciones correctas.

Recopilación de la información: El primer paso que hará nuestro sistema basado en el conocimiento es recopilar toda la información relevante que nos sea posible. Se logrará mediante un seguido de preguntas. Estas preguntas tratarán de esclarecer cuales son los gustos del lector. Gracias a las respuestas podremos hacer hipótesis y plantear evidencias.

Dentro de la recopilación de información existen diversos problemas que hemos de desengranar todavía más. Algunos como determinar la temática o la edad son sencillos, pero otros como determinar el nivel de lectura o los gustos dentro de una sub temática o sub género no lo son tan fáciles. Hemos identificado 6 de estos subproblemas derivados de la recopilación de información. Determinar la temática preferida, la edad, los elementos de la sub temática preferida, la complejidad lectora, el formato del libro, el tipo de libro en base a ventas, los libros y autores relacionados.

Análisis de la información: Conforme recopilamos la información la valoramos con tal de inferir algunos datos sobre nuestro lector y sus gustos. Algunas respuestas hacen que tomemos la decisión de hacerle unas u otras preguntas.

Valoración de la información: Una vez procesada la información de entrada, valoramos su calidad y la contraponemos con nuestras posibles soluciones. Esto hará que la calidad de nuestras soluciones aumente y/o disminuya. De esta forma obtenemos una lista de libros que posiblemente le gusten a nuestro lector. Esta lista nunca descarta un libro, es posible que con las primeras afirmaciones este libro no parezca relevante, pero que posteriormente sí lo sea.

Muestra de la solución: Una vez terminados los procesos anteriores de análisis y valoración debemos tener una lista donde se encuentran nuestras soluciones ordenadas de forma cuantitativa en base a la calidad. Se mostrarán el título de los libros de las 3 primeras soluciones junto a la calidad de esta misma.

3.3. Ejemplos del conocimiento experto extraído del dominio

Nuestro sistema toma decisiones, infiere y valora información de nuestro lector. Hemos extraído algunas de las reglas que utiliza y sus resultados. Para ello usaremos un ejemplo sencillo y directo, junto con la regla y su solución o resultado:

- Si un lector tiene una preferencia por un subgénero el sistema decidirá hacerle las preguntas pertinentes a ese subgénero.
- Si un lector indica que le gustan los best sellers automáticamente se consideran estos como parte de la solución.
- Si se indica que le gustan los libros cortos, estos ganan relevancia en la lista de soluciones
- Si se indica que tiene ebook se valora de forma positiva a los libros que se venden en formato ebook.
- Si un lector no es menor y ha indicado que le gusta la novela romántica se le pregunta si le gustan las novelas muy sensuales o no.
- Si lee mas de una hora al día y le gustan las tramas complejas, el sistema considera que es un lector avanzado.
- Dependiendo del lugar de lectura habitual transporte público, casa o descansos del trabajo, se valorará el formato del libro, tapa dura o tapa blanda y formato corto.

3.4. Descripción informal del proceso de resolución, y cómo se organizan los problemas y subproblemas

Lo primero que hace el sistema es generar una lista con todas las posibles soluciones y una valoración para cada solución que empieza en 0.

Posteriormente el sistema comienza haciendo unas preguntas generales al lector. La temática preferida, el autor, la edad, lugares de lectura,... Todas ellas con un listado de soluciones y en cada solución un número que la identifica, exceptuando los casos de si o no para las cuales deberá contestar “s” o “n”.

Algunas de estas preguntas harán que se active o no uno o mas modulos de preguntas extras. Por ejemplo si el lector es mayor de edad y ha indicado que le gustan las novelas románticas.

Una vez respondidas todas las preguntas generales empezamos con las preguntas más específicas, concernientes a la sub temática.

Durante el proceso de preguntas el sistema irá valorando las respuestas para aumentar el valor heurístico de las soluciones que cumplen dicha pregunta.

4. Formalización

4.1. Se explica detalladamente como se ha construido la ontología

La ontología se ha construido en base a los conceptos y sus características descritos en el apartado 2.1 Principalmente tenemos el concepto **Libro** y sus derivados **Libro fantasía**, **Libro novela negra**, **Libro ciencia ficción** y **Libro novela romántica**.

4.2. La ontología incluye todos los conceptos, todos los atributos, y todas las relaciones necesarias

El **Libro** posee los atributos autor, best-seller, título, edad, longitud, temática, de-moda, formato, tipo-lector y libros-similares. Este concepto representa un Libro cualquiera. Autor es una String que posee el nombre del autor que ha escrito el libro. Best-seller es un símbolo si o no, que indica si ha sido un best seller o no. Título es una String dónde se escribe el nombre del libro. Edad es un símbolo que indica la edad recomendada de lectura. Longitud es un símbolo corto o largo que indica el tamaño del libro. La temática es un símbolo que indica la temática a la que pertenece el libro. De-moda es un símbolo que puede tomar el valor de clasico o normal. El atributo formato es un símbolo cuyos valores son ebook tapa-dura tapa-blanda. El tipo-lector es un símbolo que toma los valores casual o avanzado. Por último libros-similares es un String donde hay una lista con los títulos relacionados al Libro en cuestión.

El **Libro fantasía** consta de raza y mundos-fantásticos. Raza es un símbolo en el que se identifica la raza principal que aparece en el libro, entre las cuales hay enanos, elfos, orcos, humanos y otras. mundos-fantásticos es un símbolo binario si o no. **Libro fantasía** es un elemento que a su vez es un libro y por tanto también posee los atributos de **Libro**.

El concepto **Libro ciencia ficción** posee los atributos utopía-distopía, futuros-alternativos, viajes-temporales. Utopía-distopía es un símbolo cuyo valor es utopía o distopía e indica la preferencia entre estos dos elementos. Futuros-alternativos es un símbolo binario si o no. Viajes-temporales es un símbolo binario si o no. Este elemento **Libro ciencia ficción** es, a su vez, un libro y por tanto también posee los atributos de **Libro**.

El **Libro novela negra** tiene como atributos acción y protagonistas. Los cuales hacen referencia a si hay o no acción en el libro, es un símbolo binario si o no, y que tipo de protagonistas hay en el, es un símbolo cuyos valores son

detectives, mafiosos, juzgados. Este elemento **Libro novela negra** es, a su vez, un libro y por tanto también posee los atributos de **Libro**.

El **Libro novela romántica** contiene los atributos tipo-romance y muy-sensual. El tipo-romance es un símbolo cuyos valores son histórico, contemporáneo, fantástico. Muy-sensual es un simbolo también que indica si hay o no sensualidad en el libro. Este elemento es, a su vez, un libro y por tanto también posee los atributos de **Libro**.

4.3. Documentación detallada de la ontología (se describen todos los elementos de la ontología) incluyendo un grafo de la jerarquía de conceptos

Este es el gafo de los elementos de nuestra ontología generado con protege.

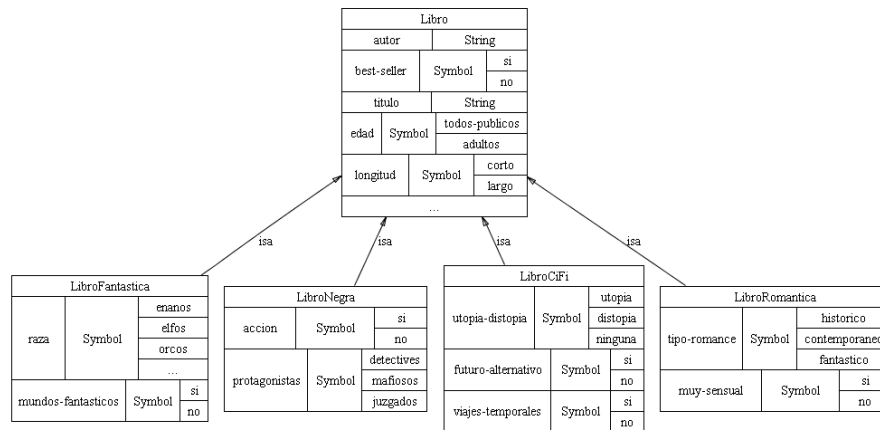


Figura 1: ontología

4.4. Se detalla el proceso de razonamiento de cada subproblema identificado

Como vemos en la ontología todos nuestros problemas tiene un elemento del cual se nutre para ser solucionado. Recordemos cuál es la lista de subproblemas que habíamos tenido en cuenta en el apartado 2.2:

1. Recopilación de la información:
 - a) Determinar la temática preferida
 - b) La edad
 - c) Los elementos de la sub temática preferida
 - d) La complejidad lectora
 - e) El formato del libro
 - f) El tipo de libro en base a ventas
 - g) Los libros y autores relacionados
2. Análisis de la información:
3. Valoración de la información:
4. Muestra de la solución:

Los puntos 2,3,4 son procesos de tratamiento de la información los cuales se encargará el SBC y para los que no necesita una representación en la ontología, para ello usará un template. En el proceso 1 toda la información que recopilamos sobre el usuario es información que debemos comparar con nuestra base de conocimiento.

Determinar la temática preferida es un subproblema trivial ya que se lo preguntamos directamente al usuario, puede escoger entres las 4 posibilidades novela romántica, novela negra, ciencia ficción y fantasía. De igual modo hacemos con la edad, que lo preguntamos directamente.

El siguiente módulo es ligeramente más complejo que los dos anteriores. Este es un recogido de preguntas relacionadas con la temática escogida en el primero de los módulos. Cada temática consta de dos a tres preguntas y hay 4 módulos que agrupan estas preguntas, uno por temática. El módulo de ciencia ficción, el de novela negra, el de novela romántica y el de fantasía. Las preguntas de estos módulos también son muy sencillas, algunas de si o no y otras de selección entre algunas de las posibilidades.

La complejidad lectora es un módulo que evalúa si el lector es un lector casual o un lector avanzado. Este módulo consta de dos preguntas, una de ellas relacionada sobre el tiempo de lectura que posee el lector y la otra en base a sus gustos en las tramas. Gracias a estas preguntas podemos derivar y saber si se trata de un lector casual o un lector avanzado.

El formato del libro es un módulo que pretende estudiar qué tipo de formato de libro es capaz de soportar y/o prefiere el lector. Este módulo consta de tres preguntas. Se le pregunta en qué lugares suele leer, que tamaño de libros prefiere y si posee o no un ebook.

El tipo de libro en base a ventas, este módulo pretende averiguar si al lector le interesan libros más o menos famosos, si le interesan best sellers o no y si prefiere un clásico o un libro que esté de moda.

Este último módulo le pregunta al lector cuál es su autor favorito y su libro favorito para posteriormente aumentar el valor de aquellas soluciones cuyo autor sea el autor favorito y las soluciones cuyas características sean las características del libro favorito del lector.

4.5. Justificación de la metodología de resolución de problemas más adecuada y como los problemas identificados en la fase anterior encajan en la metodología

En base a los apartados anteriores hemos tomado la decisión de utilizar la clasificación heurística en nuestro SBC. En clase se ha explicado que este sistema consta de 3 etapas (abstracción, asociación heurística y refinamiento)

La fase de abstracción es la fase relacionada con las preguntas. Gracias a ellas y al conocimiento que tenemos en la materia podemos hacernos una idea de como es nuestro lector, los gustos que tiene y aquello que le interesa o está buscando. Con toda esta información podemos generar un “libro abstracto” que representa el libro perfecto que nuestro usuario querría leer.

La fase de asociación heurística es en nuestro sistema el proceso que hacemos regularmente de ir valorando nuestra lista de soluciones e ir la comparando con este “libro abstracto” del que hemos hablado en el párrafo anterior. Con esta comparación buscamos los libros cuyas características son como las de este.

La fase de refinamiento es el último paso donde solo nos queda limpiar de la lista aquellos libros que no pueden ser impresos. Por ejemplo los recomendados para adultos cuando el lector es menor. Es en esta fase también donde ordenamos nuestra lista de soluciones. De este modo las soluciones con un mayor heurístico quedarán primero.

Para finalizar printamos los tres primeros libros de la lista no impresos anteriormente.

5. Implementación

5.1. Construcción de la ontología

Hemos utilizado Protégé para la generación la ontología definitiva, utilizando la clase Libro como superclase, y 4 subclases de libros (LibroCiFi, LibroFantastica, LibroRomantica, LibroNegra) que corresponden a cada una de las diferentes temáticas principales de la novela de ficción.

Sin embargo, para el versión inicial del programa utilizamos una versión simplificada de estas clases, escritas directamente en el fichero de CLIPS (*.clp). Esto nos permitió agilizar el desarrollo de la solución, y a detectar rápidamente posibles problemas.

5.2. Construcción de la base de datos

Al igual que la ontología, utilizamos Protegé para crear la versión final de las instancias de libros. Para ello creamos 8 libros de cada temática, intentando diversificar los valores para los slots -esto nos permitirá dar recomendaciones apropiadas a una amplia variedad de perfiles de lector.

Para el desarrollo de la versión inicial utilizamos otra vez una base de datos más reducida, y escrita directamente en el fichero .clp -agilizando así el desarrollo, y permitiendo la creación de instancias específicas para probar funcionalidades concretas del programa (por ejemplo el filtro de libros por edad).

5.3. Construcción del programa

5.3.1. Templates

Para representar una recomendación, hemos optado por utilizar el siguiente template:

```
(deftemplate recomendacion
  (slot libro)
  (slot rating))
```

Una recomendación es básicamente un par ¡Libro, rating! que asocia cada instancia de Libro en nuestra base de datos con una puntuación. Al iniciar la ejecución del programa, creamos todas las recomendaciones; es decir, para cada libro creamos un hecho:

```
(assert (recomendacion (libro ?libro) (rating 0)))
```

La temática preferida del usuario la hemos definido también con un template:

```
(deftemplate tematica "tematica preferida"
  (slot nombre))
```

Cada una de las 4 temáticas principales tiene su propio símbolo para el slot nombre:

- Ciencia ficción: `ciencia-ficcion`
- Novela fantástica: `fantastica`
- Novela romántica: `romantica`
- Novela negra: `negra`

5.3.2. Variables

Nuestro programa utiliza una sola variable global, que lleva la cuenta de cuántas recomendaciones hemos mostrado al lector:

```
(defglobal ?*recomendacionesImprimidas* = 0)
```

Esta variable se incrementará a medida que imprimamos los libros recomendados.

5.3.3. Hechos

Hemos comentado que las recomendaciones eran en realidad una colección de hechos en la BD, creada al iniciarse el programa.

Pero también utilizamos hechos para “dirigir” la ejecución de algunas partes del programa; por ejemplo mediante la anteriormente descrita temática:

Iniciamos la ejecución sin una temática definida:

```
(tematica (nombre desconocida))
```

Y tras preguntar al usuario cuál es su temática preferida, modificaremos este hecho con la respuesta, por ejemplo:

```
(modify ?t (nombre fantastica))
```

Este hecho nos permitirá, más adelante, decidir qué preguntas de temática debemos realizar.

En la descripción de cada módulo describiremos con detalle todos los hechos, y cómo los empleamos junto con las reglas.

5.3.4. Funciones

Tenemos básicamente 2 tipos de funciones; las de preguntas al usuario:

ask-question: escribe la pregunta, y permite al usuario dar la respuesta de entre una serie de opciones. Retorna un símbolo que será la opción seleccionada.

si-o-no-p: particularización de la anterior función -la respuesta será sí o no, y retorna un valor booleano.

pregunta-numerica: Realiza una pregunta a la que el usuario debe responder con un número en el intervalo especificado.

El resto de funciones permiten modificar el rating (o puntuación) de las recomendaciones:

incrementa-rating-recomendacion: dada una recomendación y un incremento, suma ese valor al rating de la recomendación.

incrementa-rating-libros: dado un slot, un valor, y un incremento, suma ese incremento al rating de todas las recomendaciones cuyo libro tiene ese slot con el valor indicado. El slot debe ser uno de los de la clase Libro.

incrementa-rating-libros-lista: esta función igual que la anterior, pero para multislots; aumentará el rating de las recomendaciones cuyo libro tenga en el multislot al menos un valor idéntico al especificado.

incrementa-rating-libros-tematica: Como cada una de las subclases de Libro tiene unos slots particulares, hemos creado una función que permite incrementar el rating de las recomendaciones de libros de una temática concreta.

5.3.5. Módulos

Como dijimos anteriormente, hemos creado 9 módulos distintos. Pasemos a explicar cada uno en detalle.

1 - MAIN

En este módulo inicializamos todas las recomendaciones, y establecemos la temática preferida del usuario como desconocida.

2 - determina-tematica

En este módulo dejamos que el usuario seleccione su temática favorita. Sin importar cuál sea su respuesta, debemos hacer 2 cosas:

Establecer el hecho (tematica (nombre ?)) Modificar la puntuación de todos los libros de esa temática, de forma que las recomendaciones sólo puedan ser de libros de la temática. Esto lo conseguimos sumando 1000 -un valor superior a la suma del resto de incrementos de rating del programa.

3 - preguntas-edad

Simplemente preguntamos la edad al usuario, y establecemos el hecho lector-menor si o lector-menor no. Esta pregunta no tiene influencia alguna en el rating de los recomendaciones -su finalidad es tan solo ocultar o no ciertos libros.

4 - preguntas-tematica

A continuación hacemos al usuario las preguntas específicas de temática. Estas aportan un incremento pequeño (de 1, concretamente), ya que pensamos que las características afectadas no son tan importantes.

Ciencia ficción:

- ¿Te gustan las novelas de futuros alternativos? (s/n)
- ¿Te gustan los viajes en el tiempo? (s/n)
- ¿Utopía o distopía? (utopía, distopía, ninguna)

Fantástica:

- ¿Te gustan los mundos fantásticos? (s/n)
- ¿Cuál de estas razas prefieres? (enanos, elfos, orcos, humanos, otros)

Romántica:

- ¿Qué tipo de romance prefieres? (histórico, contemporáneo, fantástico)
- ¿Te gustan las novelas románticas muy sensuales? (s/n)

Negra:

- ¿Te gusta que haya mucha acción? (s/n)
- ¿Qué tipo de historia prefieres? (detectives, mafiosos, juzgados)

Todas incrementan el rating de los libros con atributos coincidentes.

Cabe notar que si el usuario es menor, la pregunta ¿Te gustan las novelas románticas muy sensuales? no le aparecerá, ya que las novelas muy sensuales son sólo para adultos.

5 - preguntas-tiempo-complejidad

Con este módulo comienzan preguntas de carácter más general. Al igual que las preguntas de temática, el incremento aportado por estas debe ser pequeño (de nuevo es de 1).

Primero intentaremos descubrir si es un lector casual o avanzado. Para ello, preguntaremos cuánto puede leer al día; si no dispone ni de una hora,

daremos por sentado que es lector casual: incrementamos el rating de los libros con ese tipo-lector.

Si dispone de más tiempo de lectura al día, le preguntamos explícitamente si le gustan las tramas complejas, ya que aunque sea un ávido lector, tal vez prefiere un tipo de lectura más ligero.

6 - preguntas-formato-longitud

Seguidamente, queremos descubrir el formato (ebook, tapa blanda o tapa dura) y longitud (corto o largo) de libro preferidos. Según dónde lea el usuario:

- si lee en transporte publico, tapa blanda, longitud libro indeterminada
- si lee en casa, tapa dura, longitud libro indeterminada
- si lee en el trabajo, tapa blanda, longitud libro corta

Y si aún no estamos seguros de si prefiere libros largos o cortos, se lo preguntamos explícitamente.

También preguntaremos explícitamente si el usuario dispone de ebook.

7 - preguntas-ventas-categoria

Las preguntas de este módulo afectan directamente al rating de las recomendaciones de:

- Best-sellers: damos más peso a los super-ventas, si es que el usuario dice que le gustan. Si no le gustan, no modificamos las recomendaciones, ya que asumimos que simplemente no le importan las ventas.
- ¿Clásico, de moda, o indiferente? Pensamos que esta clasificación puede ser relevante para algunos lectores, por lo que aumentaremos el rating de los libros correspondientes.

8 - preguntas-autor-libro-favoritos

En este módulo intentamos descubrir qué autor y obra de referencia tiene este usuario. Creemos que este tipo de información es muy útil para procurar buenas recomendaciones: simplemente por el hecho que, en la mayoría de casos, hay cierta uniformidad entre los libros de un mismo autor, o las obras similares, y esto es muchas veces garantía de éxito. Decidimos dar un incremento de 3 a las recomendaciones afectadas.

En cuanto al autor: si el usuario tiene un autor preferido, aumentamos el rating de las recomendaciones de los libros de ese autor.

En cuanto a las obras similares: si el lector tiene un libro preferido, damos más peso a los libros similares -es decir, aquellos que tengan en su lista libros-similares el título entrado por el usuario.

9 - imprime-recomendaciones

Llegado este módulo, sólo nos queda imprimir los libros de las recomendaciones que más puntuación hayan obtenido. La primera regla del módulo establece el hecho `print-sorted`, que dispara la selección de libros por imprimir.

La regla `assert-unprinted` marca todo libro como “no impreso”, vigilando de ignorar los libros para adultos si el lector resulta ser menor de edad.

La siguiente regla, `retract-print-sorted`, elimina el hecho `print-sorted`, disparando así la impresión de los 3 libros mejor recomendados.

Y de esto se encarga la regla `print-greatest`, que repetirá el siguiente proceso hasta que la global `?*recomendacionesImprimidas*` valga 3:

Encuentra la recomendación con libro “no impreso”, y rating mayor que todas las demás recomendaciones con libro “no impreso”. Imprime ese libro por el canal estándar de salida. Suma 1 a la global `?*recomendacionesImprimi-`

das*.

5.4. Versiones

A continuación describimos las 4 etapas por las que hemos pasado para desarrollar el Sistema Experto:

5.4.1. Inicial

En esta primera versión, utilizamos una ontología e instancias simplificadas, codificadas directamente en el fichero de CLIPS. Además, partimos solamente con la clase Libro, por lo que las preguntas que se realizaban eran solamente las de carácter general.

Los incrementos de rating eran constantes, ya que sólo estábamos interesados en ver si el programa funcionaba de la forma esperada. Esta fue nuestra toma de contacto con CLIPS.

5.4.2. Aumentado 1

Al igual que en la anterior versión, utilizamos una ontología e instancias simplificadas, pero esta vez añadimos las subclases LibroCiFi, LibroFantastica, LibroRomantica, LibroNegra, y las preguntas especiales para cada temática.

También añadimos aquí el filtro de libros por edad (con el hecho lector-menor), asegurándonos que su funcionamiento es correcto.

5.4.3. Aumentado 2

Por fin integramos la ontología y las instancias creadas con Protegé. Esta parte era especialmente delicada, ya que el modelo completo tenía que casar con la versión simplificada, y no romper el correcto funcionamiento del programa.

5.4.4. Definitivo

Con todo lo de Protegé integrado, ya sólo faltaba determinar qué reglas y atributos debían dar más peso a las recomendaciones correspondientes.

Optamos por dar mucho peso a la selección de temática, de forma que las recomendaciones siempre serán de libros de la temática preferida del lector.

El autor y libro preferidos del lector son lo siguiente en cuanto a peso, ya que recomendar otros libros del mismo autor, y obras similares suele dar muy buenos resultados.

El resto de preguntas aportan un incremento mínimo -sirven para refinar la selección de recomendaciones.

6. Pruebas

6.1. Prueba 1

6.1.1. Introduccion

El primer sujeto de pruebas representa un lector avanzado promedio de literatura fantástica, fan de Ron Coletti, cuyo libro favorito es uno de los de su obra llamado “Sistema experto”. Dado a que es un lector avanzado, está

acostumbrado a la complejidad, libros largos y a leer varias horas al día, con preferencias variadas sobre el género ya que, al ser un lector experimentado, las pequeñas diferencias de cada novela resultan irrelevantes para él.

Esta prueba ha sido determinada principalmente por que al ser la primera, queremos obtener el resultado más predecible posible para comprobar que el sistema de recomendación diseñado es medianamente adecuado.

6.1.2. Resultado Esperado

Cada variable asignada por el usuario, a parte de la temática que puntúa de sobremanera a las demás, cuenta igual, por eso, no es muy difícil predecir que tipo de libro debería aparecer, puesto que la mayor puntuación vendrá determinado por el número de variables en común con el parámetro de entrada.

Así pues, podemos esperar que en los resultados incluyan: -La propia obra de Ron Colletti, “Sistema Experto, u otras obras de Ron Coletti. -Obras similares a obras de Ron Colletti, especialmente similares a “Sistema Experto” -Obras más parecidas a las pequeñas preferencias del usuario.

6.1.3. Entrada CLIPS

–Sistema Experto de Recomendación de Libros–

¿Que tematica estás buscando? 1 - Fantastica 2 - Ciencia ficcion 3 - Novela romantica 4 - Novela negra

1

¿Que edad tienes? [1, 100]

33

¿Te gustan los mundos fantásticos? (s/n)

s

Cual de estas razas prefieres? 1 - Enanos 2 - Elfos 3 - Orcos 4 - Humanos

5 - Otras

3

¿Lees mas de una hora al dia? (s/n)

s

¿Te gustan las tramas complejas? (s/n)

s

¿Dónde sueles leer? 1 - Transporte publico 2 - Casa 3 - Descansos del trabajo

2

¿Prefieres los libros largos? (s/n)

s

¿Tienes ebook? (s/n)

n

¿Te gustan los best-sellers? (s/n)

s

¿Que tipo de libro te gusta? 1 - Clásico 2 - De moda 3 - Indiferente

3

¿Tienes un autor preferido? (s/n)

s

¿Quién es tu autor preferido?

Ron Coletti

¿Tienes un libro preferido? (s/n)

s

¿Cual es tu libro preferido?

Sistema Experto

6.1.4. Sortida CLIPS

#1 - Depredador, de Ron Coletti

#2 - Sistema Experto, de Ron Coletti

#3 - El señor de los pardillos, de Jeremias Sauron

6.1.5. Resultado

Tras comparar los resultados ejecutados por CLIPS y los previstos por nosotros basándonos en la teórica ejecución del programa, podemos afirmar que, en casos básicos como este el funcionamiento es aceptable: muestra “Depredador” como principal recomendación, ya que es de la temática y autor preferidos del usuario, y además cumple otras preferencias explícitas del usuario (best-seller, mundos fantásticos). “Sistema Experto” es la siguiente recomendación, ya que es del mismo autor y también atributos buscados por el usuario. Y finalmente, “El señor de los pardillos”, que reúne casi todas las características para

gustar al lector -tanto las explícitas (mundos fantásticos, orcos, libro largo), como las deducidas (formato de tapa dura, libro adulto).

6.2. Prueba 2

6.2.1. Introduccion

Para este caso utilizaremos una estructura similar a la del caso 1, pero esta vez, para asegurarnos que la distinción de temáticas funcione correctamente, cambiaremos la temática para Ciencia-Ficción, así pues, nuestro usuario tendrá exactamente las mismas preferencias globales, solo que para autor y libro favorito, estos sean 2 de la rama de la ciencia ficción.

Por ende el autor favorito será “Joseph Wolf” con su obra llamada “BrazeRunner”.

6.2.2. Resultado Esperado

Como hemos dicho puesto que la entrada es una versión similar a la de la prueba 1, podemos esperar un resultado que presente similitudes con los resultados de la prueba 1, pero en una versión de libros de Ciencia Ficción.

6.2.3. Entrada CLIPS

Sistema Experto de Recomendación de Libros

¿Qué temática estás buscando? 1 - Fantástica 2 - Ciencia ficción 3 -
Novela romántica 4 - Novela negra

2

¿Qué edad tienes? [1, 100]

33

¿Te gustan los viajes en el tiempo? (s/n)

n

¿Utopía o distopía? 1 - Utopía 2 - Distopía 3 - Ninguna

1

¿Te gustan las novelas de futuros alternativos? (s/n)

s

¿Lees más de una hora al día? (s/n)

s

¿Te gustan las tramas complejas? (s/n)

s

¿Dónde sueles leer? 1 - Transporte público 2 - Casa 3 - Descansos del
trabajo

2

¿Prefieres los libros largos? (s/n)

s

¿Tienes ebook? (s/n)

n

¿Te gustan los best-sellers? (s/n)

s

¿Qué tipo de libro te gusta? 1 - Clásico 2 - De moda 3 - Indiferente

3

¿Tienes un autor preferido? (s/n)

s

¿Quién es tu autor preferido?

Joseph Wolf

¿Tienes un libro preferido? (s/n)

s

¿Cuál es tu libro preferido?

BrazeRunner

6.2.4. Salida CLIPS

#1 - BrazeRunner, de Joseph Wolf

#2 - Hemoplastia, de Tom Anderson

#3 - La granja, de Joseph Wolf

6.2.5. Resultado

En este caso, como las variables de las preferencias si coinciden con las del libro favorito en cuestión por lo que el sistema puede recomendar “BrazeRunner” con facilidad, asi como otras obras que comparten las mismas características, que a diferencia del caso 1, son más parecidas a las preferencias del libro deseado del usuario.

6.3. Prueba 3

6.3.1. Introduccion

Para la siguiente prueba utilizaremos un caso de usuario menos experimentado. El usuario representa un menor de edad que busca una lectura más o menos simple, casi sin tener ninguna experiencia como lector, i no poder determinar un autor o libro favorito.

Nos Centraremos en la temática de Literatura de Ciencia Ficción, puesto que esta presenta más ejemplares de géneros para todos los públicos.

6.3.2. Resultado Esperado

Al no disponer de escritores y libros favoritos es difícil suponer un resultado concreto, ya que en general cada ejemplar es diferente el uno del otro en términos de preferencias como para suponer un resultado a partir de unos valores de entrada sin datos concretos como Autor o libro favorito. Lo que si podemos asegurar es que el resultado debería ser un seguido de libros de Ciencia-Ficción, probablemente para lectores casuales.

6.3.3. Entrada CLIPS

–Sistema Experto de Recomendación de Libros–

¿Que temática estás buscando? 1 - Fantástica 2 - Ciencia ficción 3 -
Novela romántica 4 - Novela negra

2

¿Qué edad tienes? [1, 100]

12

¿Te gustan los viajes en el tiempo? (s/n)

s

¿Utopía o distopía? 1 - Utopía 2 - Distopía 3 - Ninguna

3

¿Te gustan las novelas de futuros alternativos? (s/n)

s

¿Lees más de una hora al día? (s/n)

n

¿Dónde sueles leer? 1 - Transporte público 2 - Casa 3 - Descansos del trabajo

2

¿Prefieres los libros largos? (s/n)

n

¿Tienes ebook? (s/n)

s

¿Te gustan los best-sellers? (s/n)

n

¿Qué tipo de libro te gusta? 1 - Clásico 2 - De moda 3 - Indiferente

2

¿Tienes un autor preferido? (s/n)

n

¿Tienes un libro preferido? (s/n)

n

6.3.4. Salida CLIPS

#1 - 1987, de Joseph Wolf tiene puntuacion = 1004

#2 - La granja, de Joseph Wolf tiene puntuacion = 1003

#3 - Un dia como Vulcano, de Leonardo Nimoyo tiene puntuacion =
1003

6.3.5. Resultado

Las recomendaciones se aproximan a los resultados dichos previamente.

6.4. Prueba 4

6.4.1. Introduccion

Para la cuarta prueba nos centraremos en los libros de novela romántica. Utilizaremos un usuario mayor de edad para que no se limite las preguntas y intentaremos crear un punto intermedio entre los casos 2 y 3, así pues la idea es crear un lector casual, que tenga un mínimo de conocimiento en ese ámbito de novelas, ponemos como escritor favorito a Susan Sander, quien tiene diversas entradas en esta temática.

6.4.2. Resultat Esperado

Es esperable que haya mas de un titulo de Susan Sander entre las recomendaciones, asi como todos los titulos que mantengan mas similitudes.

6.4.3. Entrada CLIPS

–Sistema Experto de Recomendación de Libros–

¿Qué temática estas buscando? 1 - Fantástica 2 - Ciencia ficción 3 -
Novela romántica 4 - Novela negra

3

Fin pregunta tematica

¿Qué edad tienes? [1, 100]

20

Fin pregunta edad

¿Te gustan las novelas románticas muy sensuales? (s/n)

s

¿Qué tipo de romance prefieres? 1 - Histórico 2 - Contemporáneo 3 -
Fantástico

2

Fin preguntas temática

¿Lees más de una hora al día? (s/n)

n

¿Dónde sueles leer? 1 - Transporte público 2 - Casa 3 - Descansos del trabajo

1

¿Prefieres los libros largos? (s/n)

n

¿Tienes ebook? (s/n)

s

¿Te gustan los best-sellers? (s/n)

s

¿Qué tipo de libro te gusta? 1 - Clásico 2 - De moda 3 - Indiferente

2

¿Tienes un autor preferido? (s/n)

s

¿Quién es tu autor preferido?

Susan Sander

¿Tienes un libro preferido? (s/n)

n

6.4.4. Salida CLIPS

#1 - El triangulo del amor, de Susan Sander tiene puntuacion = 1007

#2 - Pasion de Centollos, de Susan Sander tiene puntuacion = 1007

#3 - Ojala yo tuviera este primero, de Susan Sander tiene puntuacion
= 1005

6.4.5. Resultado

Com era esperable, las entradas de Susan Sander predominan en las recomendaciones, puede resultar sorprendente no encontrar títulos de otros escritores, sin embargo el motivo es simple. Al no especificar un libro preferido, no puede haber libros similares de otros autores que se asemejen a él, por ende solo los libros de Susan Sander parten con una predilección que ninguno de los demas tiene, además de las posibles preferencias que se puedan encontrar en dichos libros.

6.5. Prueba 5

6.5.1. Introduccion

Para el último caso tocaremos un caso muy concreto de novela Negra, con el fin de ver resultados exactos sobre un caso concreto, asi pues interpretamos como objetivo el usuario promedio que la base de datos espera para un libro concreto de Novela negra, con el fin de que salga este, y sus dos libros más parecidos.

6.5.2. Resultado Esperado

El libro elegido para probar la eficacia del sistema sera “Spanish History Z” de Ramon Cajon, con el fin de que el sistema devuelva su secuela “Spanish History Y” y su similitud literaria alternativa ajena a Ramon Cajon,

.En busca de la ontología perdida”.

Puesto que nuestro objetivo mínimo es sacar como recomendación el libro elegido, nuestro usuario será un lector casual adulto que busca una novela negra de longitud corta en formato de e-book.

6.5.3. Entrada CLIPS

–Sistema Experto de Recomendación de Libros–

¿Qué temática estás buscando? 1 - Fantástica 2 - Ciencia ficción 3 -
Novela romántica 4 - Novela negra

4

Fin pregunta tematica

¿Qué edad tienes? [1, 100]

55

Fin pregunta edad

¿Te gusta que haya mucha acción? (s/n)

s

¿Qué tipo de historia prefieres? 1 - Detectives 2 - Mafiosos 3 - Juzgados

2

Fin preguntas temática

¿Lees más de una hora al día? (s/n)

n

¿Dónde sueles leer? 1 - Transporte público 2 - Casa 3 - Descansos del trabajo

3

¿Tienes ebook? (s/n)

s

¿Te gustan los best-sellers? (s/n)

n

¿Qué tipo de libro te gusta? 1 - Clásico 2 - De moda 3 - Indiferente

3

¿Tienes un autor preferido? (s/n)

s

¿Quién es tu autor preferido?

Ramon Cajon

¿Tienes un libro preferido? (s/n)

s

¿Cuál es tu libro preferido?

Spanish History Z

6.5.4. Salida CLIPS

#1 - Spanish History Y, de Ramon Cajon

#2 - Spanish History Z, de Ramon Cajon

#3 - En busca de la ontología perdida, de Esteban Reyes

6.5.5. Resultado

Tal y como esperabamos los resultados coinciden con los parametros objetivos, por lo que se confirma el uso correcto del uso de variables de entrada.

7. Conclusión y mejoras

Hemos creado un Sistema Basado en el Conocimiento de recomendación de libros: a partir de información acerca del lector, seleccionamos libros que serán probablemente de su agrado.

Para desarrollar el proyecto hemos aplicado la metodología en cascada, utilizando Protegé para crear la ontología y las instancias de los libros, y CLIPS para el Sistema Experto.

Hemos visto que las recomendaciones procuradas son las apropiadas si el usuario ha contestado las preguntas con sinceridad. En realidad, cuanto más “coherente” es el usuario (por ejemplo, si a un usuario le gustan los libros de moda, en particular le gustarán los best-seller), mejor funciona el sistema.

A pesar de todo, nos dimos cuenta de que, utilizando una base datos más extensa, y basada en obras reales, el sistema podría dar mayor variedad recomendaciones perfiles muy parecidos.