

Documentació de la pràctica de sistemes basats en el coneixement



Federico Rubinstein Pérez,
Roger González Herrera,
Luis Oriol Soler Cruz

FIB • UPC

Quadrimestre Tardor - Curs 2019/2020

INDEX

1. El problema	3
1.1. Viabilitat de construir un SBC	4
1.2. Fonts del coneixement	4
1.3. Objectius	5
2. Conceptualització	6
2.1. Conceptualització del problema	6
2.1.1. Creació del conjunt de llibres disponibles.	6
2.1.2. Càlcul de la dificultat d'un llibre.	6
2.1.3. Càlcul de la valoració d'un llibre.	7
2.1.4. Gestió de la informació.	7
2.2. Elements del domini	9
2.2.1. Característiques del Llibre.	9
2.2.3. Característiques d'un autor	9
2.3. Divisió en subproblemas	9
2.3.1. Recollida de dades de l'usuari.	10
2.3.2. Anàlisi de les preferències de l'usuari.	10
2.3.3. Presentació de la solució.	10
2.4. Procés de la solució.	10
2.5. Organització de la solució.	11
3. Formalització	12
3.1. Desenvolupament de l'ontologia	12
3.1.1. Definició de les classes	12
3.2. Mètode de resolució	14
3.2.1. Abstracció de les dades	14
3.2.2. Associació heurística	14
3.2.3. Transformació a una solució tangible	14
4. Implementació	15
4.1. Mòduls	15
4.2. Prototip	15
5. Jocs de prova	16
6. Conclusions	21

1. El problema

Aquesta pràctica està dissenyada per a ajudar a entendre els diferents models de Intel·ligència Artificial aplicables per a solucionar un problema que necessita de coneixement expert per a poder obtenir una solució.

El problema que proposa la pràctica és el de desenvolupar una eina de recomanació de llibres per a un usuari concret. Aquesta recomanació ha de ser el més proper possible a la resposta que ens donaria el nostre llibreter de confiança, ja que aquesta resposta estarà molt més dirigida que la d'un simple sistema de recomanacions que no té en compte informació sobre el lector.

És per aquest motiu que hem d'esbrinar com actua un llibreter a l'hora de recomanar un llibre i simular aquest comportament. Per tant, el nostre sistema basarà les seves recomanacions en diverses característiques de l'usuari i en el tipus de llibre que està buscant aquest.

Per tant, tenim diversos tipus de dades:

- Dades del lector: recopilació de dades bàsiques de l'usuari, com per exemple l'edat, amb quina freqüència llegeix, on llegeix, etc. Aquestes dades són útils per a assignar algunes recomanacions per defecte a l'usuari, independentment de les seves preferències.
- Preferències de l'usuari: recullen els gustos de l'usuari, és a dir, els gèneres, autors i temàtiques preferits. El sistema utilitzarà aquesta informació per a saber quin tipus de llibre s'està buscant.
- Llibres llegits: és una llista de llibres que l'usuari ja ha llegit, per tal d'excloure-los de les recomanacions finals.

Amb aquesta informació ja hauriem de ser capaços de complir l'objectiu de la pràctica, és a dir, escollir els 3 llibres que segurament l'interessin més a l'usuari. Tot i així, en el cas de que tota aquesta informació no sigui suficient, sempre podem recomanar al lector els llibres més populars que s'acostin a les dades donades.

1.1. Viabilitat de construir un SBC

El problema que ha de resoldre el sistema es el de trobar un conjunt de fins a 3 llibres per a recomanar a l'usuari. És un problema de cerca en el qual hem d'escollir, entre un conjunt gran de llibres, els que més s'aproximin als requisits de l'usuari.

Aquesta cerca és molt costosa i complexa com per a solucionar-la de forma algorítmica.

Per tant, el nostre problema és millor solucionar-lo amb un Sistema Basat en el Coneixement, a causa de la falta de criteris per a endreçar el contingut i la dificultat de trobar una estratègia per a recórrer el contingut de manera eficient.

Si podem obtenir el coneixement necessari per a representar el perfil de l'usuari i les característiques dels llibres (és a dir, coneixement específic del domini), podem aplicar les regles i restriccions d'un SBC per a reduir la complexitat i els costos de la cerca a algo raonable d'utilitzar.

A més a més, veiem que podem dissenyar un sistema que compleixi totes les característiques típiques dels Sistemes Basats en el Coneixement i que el diferenciïn dels sistemes de software tradicionals.

Podem dissenyar un sistema flexible (ja que el problema que es pot representar al domini pot ser variat), que simuli un comportament racional (com hem dit, per exemple, que simuli el procés que faria el llibreter de confiança), que operi en un entorn ric en informació (es pot aconseguir molta informació i molt diversa sobre el domini), que utilitzi informació simbólica i llenguatge natural i que se li pugui afegir la capacitat d'aprendre.

1.2. Fonts del coneixement

Les fonts del coneixement són aquelles fonts que s'utilitzen en un SBC per a conseguir la informació necessària per a resoldre el problema.

És important buscar les fonts del coneixement que s'utilitzaran a l'hora d'abordar el problema.

En el nostre cas, les fonts del coneixement eren:

- Consultar en internet per a construir una base de dades de llibres suficientment gran amb camps útils per a cada llibre. Els camps de cada llibre que hem escollit per al nostre sistema són: títol, autor, any, dificultat, gènere, si és best seller, si és per a joves, temàtica, longitud i valoració. Hem escollit llibres només dins del domini de les

noveles de ficció, ja que sinó la base de dades seria massa gran i no podríem centrar-nos tant.

- L'altre font de coneixement de la qual disposa el nostre sistema és l'usuari. Quan l'usuari executa el sistema se li fan una sèrie de preguntes que serveixen per a conèixer-lo millor i per a saber les seves preferències. És gràcies a les respostes a aquestes preguntes que podem recomanar-li llibres específics. Aquestes preguntes les hem escollit gràcies a la informació recopilada en la cerca per internet i les que veiem més adients per a solucionar el problema inicial.

1.3. Objectius

L'objectiu d'aquesta pràctica és el de enfrontar-se a un problema que necessita de coneixement expert per a poder solucionar-lo. El sistema plantejat ha de resoldre tasques d'anàlisi, identificació i construcció d'una solució.

L'objectiu del nostre sistema és resoldre les següents tasques:

- Anàlisi d'una base de dades de llibres que li hem proporcionat.
- Interacció amb l'usuari per tal d'obtenir tota la informació necessària sobre l'usuari, les seves preferències i el conjunt de llibres de la nostra base de dades que ja ha llegit.
- Anàlisi de les dades obtingudes de l'usuari per tal de recomanar els llibres adients, aprofintant l'aprenentatge que tenim dels llibres del sistema.
- Endreçar els llibres de manera que els que segurament tinguin més interès per a l'usuari quedin a l'inici. Deixem els llibres que l'usuari ja ha llegit al final, per a que només apareguin en cas que s'hagi llegit tots els llibres de la base de dades.
- Mostrar a l'usuari els 3 llibres amb més possibilitats de que li interessin. També mostrem la informació d'aquest llibre i perquè ha sigut escollit.

2. Conceptualització

Hem de dissenyar una estructura de dades que ens permeti guardar llibres i les seves relacions, per a que posteriorment puguem decidir utilitzant el sistema basat en coneixement quins llibres s'adapten millor al perfil obtingut de l'usuari, és a dir del lector.

Per poder fer això ens hem estructurat en crear el perfil del lector, al qual li preguntem l'edat, la seva nacionalitat, la freqüència de lectura, el lloc on li agrada llegir i si es deixa guiar per els llibres populars (bestseller) o per les valoracions. Després creem també un perfil del conjunt de llibres que pot agradar al lector, és a dir el gènere, la temàtica els autors preferits i els llibres que ja ha llegit.

Cada atribut obtingut té un impacte a la puntuació de cada llibre, aquest impacte pot ser diferent a la d'un altre atribut o igual, i es a partir d'aquesta puntuació que valorem quin és el o els llibres adequats per recomanar al lector.

2.1. Conceptualització del problema

2.1.1. Creació del conjunt de llibres disponibles.

Per formar la base de dades que conté tots els llibres hem extret informació de les webs recomanades a l'enunciat de la pràctica, a més hem fet servir diferents webs que valoraven llibres per saber altres atributs. Al final hem aconseguit un total de 50 llibres que fan que la nostra pràctica tingui suficient espai per generar solucions variades i per a diferents perfils de lectors fagi diferents recomanacions.

2.1.2. Càlcul de la dificultat d'un llibre.

Per a calcular la dificultat d'un llibre hem tingut en compte la mida d'aquest llibre, en pàgines, l'any d'edició i també els millors comentaris que feien els usuaris de les pàgines recomanades per l'enunciat, a partir d'allà generem un valor entre 1 i 10 que es orientatiu i que en cap cas busca ser un valor exacte per cada llibre, ja que es un factor difícil de mesurar.

2.1.3. Càlcul de la valoració d'un llibre.

Per fer el càlcul de la valoració hem buscat a una web especialitzada en valoracions de llibres, https://www.goodreads.com/review/recent_reviews, els usuaris d'aquesta pàgina valoren el llibre i aporten informació addicional com el gènere, comentaris sobre el llibre i un resum del llibre. Com que la valoració és una puntuació entre 0 i 5 i nosaltres la vam establir de 1 a 10 hem multiplicat la nota per dos i l'hem arrodonit perquè sigui un valor enter. Vam escollir aquesta pàgina perquè el fet de buscar-ho en pàgines on ho pots comprar es barreja el factor de valoració del llibre com a producte de compra i el que nosaltres volíem era que no influís el preu, que simplement es valori el llibre com a un producte per llegir i no com a un objecte amb valor.

2.1.4. Gestió de la informació.

Potser el dilema més gran que vam tenir era com enfocar el recull de dades de l'usuari (és a dir, el lector) i com això podia influir al resultat. Hem d'escollir en que ens fixem a l'hora de preguntar què volem saber i com això influeix en la tria que farà el nostre sistema a l'hora de recomanar.

Per fer-ho vam generar preguntes per aconseguir informació de l'usuari. Les preguntes eren les següents:

- *Qué edad tiene?*

Aquesta pregunta la fem per valorar els llibres que son de caire Juvenil o no.

- *Con qué frecuencia lee?*

Aquesta pregunta és tancada i la resposta pot ser: Diàriament, Habitualment, Esporàdicament o Mai. L'objectiu és valorar factors com si és factible que pugui llegir un llibre de moltes pàgines o si és millor un llibre fàcil de llegir.

- *En qué lugar suele leer?*

Aquesta pregunta també és tancada i les seves respostes poden ser Casa, Biblioteca, Aire Lliure, Transport públic o Indiferent. Volem valorar si es factible que l'usuari pugui llegir un llibre llarg o com assignar un llibre de dificultat baixa o alta.

- *Le gustan los libros Populares (bestseller)?*

Amb aquesta pregunta valorem els llibres que son populars i si agraden o no al usuari.

- *A la hora de escoger un libro se deja llevar por las valoraciones?*

Amb aquesta pregunta valorem fins a quin punt hem de donar importància a les valoracions.

- *Prefiere autores extranjeros?*

Amb aquesta pregunta valorem la importància d'un llibre escrit per un autor estranger.

- *De qué nacionalidad eres?*

Gràcies a la pregunta anterior valorem positivament els llibres escrits per autors estrangers o no al seu país, depenent de com ha contestat a la pregunta anterior.

Fins aquí hem obtingut la informació purament del l'usuari. Les següents preguntes son les que van més enfocades al gust que pot tenir el lector i a com acabem de definir millor els llibres que més li correspon.

- *Escoja sus generos preferidos*

Amb aquesta pregunta valorem els gèneres que li agraden al lector.

- *Escoja sus temáticas preferidas*

Amb aquesta pregunta valorem les temàtiques que son preferibles per al lector.

- *Escoja sus autores preferidos*

Amb aquesta pregunta valorem els autors que poden ser més adients per el nostre lector.

Fins aquí ja podem obtenir una sèrie de llibres que s'acomoden a la necessitat del lector i per tant es una bona recomanació. Nosaltres vam voler anar més enllà i tenir en compte que possiblement arribat a aquest punt ens podia sortir a la solució un llibre que ja hagués llegit el lector i per tant vam escollir fer una última pregunta:

- *Escoja los libros que ha leído*

Amb aquesta pregunta descartem possibles llibres que l'usuari ja ha llegit, així la recomanació és més afinada.

2.2. Elements del domini

A continuació llistem els diferents elements del domini i els seus camps.

2.2.1. Característiques del Llibre

- Nom
- Número de pàgines
- És Bestseller
- És Juvenil
- Dificultat
- Valoració
- Autor
- Genere
- Temàtica

2.2.2. Característiques de l'Usuari

- Edat
- Freqüència de lectura
- Lloc de lectura
- Nacionalitat
- Li agrada els BestSeller
- Es guia per les recomanacions
- Li agraden autors estrangers

2.2.3. Característiques d'un autor

- Nom
- Nacionalitat

2.3. Divisió en subproblemas

A continuació fem una breu descripció dels diversos subproblemes que ha de resoldre el nostre sistema per tal de solucionar la tasca principal.

2.3.1. Recollida de dades de l'usuari.

Tal i com hem esmentat anteriorment, al iniciar el programa, l'usuari respon a una sèrie de preguntes que fan referència a característiques pròpies del lector. Un cop fet aquest pas, ha de respondre la segona part de preguntes que fan referència a preferències sobre llibres que son disponibles a la nostra Ontologia.

2.3.2 Anàlisi de les preferències de l'usuari.

El programa es dedica a analitzar les dades recollides sobre l'usuari. La forma en que ho hem plantejat és amb un camp de puntuació. Segons les preferències de l'usuari apliquem segons cada camp una puntuació a cada llibre de manera que alguns atributs poden sumar més que d'altres i així obtenim una classificació dels llibres més adequats per cada usuari.

2.3.3. Presentació de la solució.

Un cop feta la classificació dels llibres, i ordenats de manera que el més adient, és a dir, el de més puntuació, sigui el primer, ja tenim les propostes que el nostre recomanador pot fer.

En el cas de la nostra pràctica el resultat seran els 3 llibres més puntuats, és a dir, els 3 primers.

2.4. Procés de la solució.

Hem assignat un valor a cada un dels aspectes segons la seva importància, aquest valor és el que sumarem a la puntuació en el cas que el factor que hem preguntat correspongui amb les necessitats del lector. Ens basem en les preferències per ajustar la puntuació de cada llibre al que l'usuari espera, i aïllem els paràmetres que no comporten cap impacte a la puntuació. Així doncs, després de recollir totes les dades necessàries, ja podem aplicar aquests mètodes i esperar a que el programa ens generi la solució final.

2.5. Organització de la solució.

Partim de la informació que ens ha proporcionat l'usuari i de que sabem que és un problema que es pot resoldre amb un Sistema Basat en Coneixement. Els passos que hem de seguir són l'abstracció de les dades, l'obtenció d'aquestes i la associació heurística amb els resultats; tot això ens genera una solució final que nosaltres la interpretem de manera que mostrem els 3 primers llibres.

3. Formalització

En aquest apartat veurem una visió del problema des del punt de vista de l'enginyer del coneixement, en el nostre cas, el llibreter de confiança. En aquesta etapa decidirem quins mecanismes de entre els que utilitza l'expert en el problema és més adequat per a les nostres circumstàncies.

3.1. Desenvolupament de l'ontologia

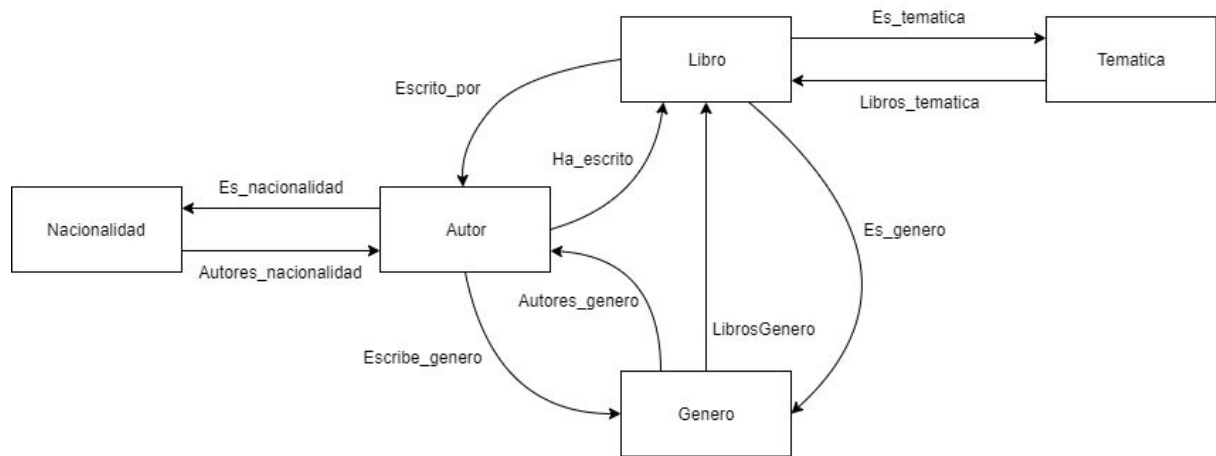
Un cop definits els conceptes del domini, podem traslladar aquests conceptes per a fer la ontologia. Aquesta ha de representar l'usuari, els llibres i les seves característiques; és a dir els autors i les seves nacionalitats, els gèneres i les temàtiques.

El terme més important i on ens hem de centrar de la ontologia són els llibres, perquè són aquells l'objecte que ha de ser recomanat i per tant on hem centrat molta part de l'estudi. També, però, és centrar-nos en l'usuari i saber exactament i afinament quines són les seves preferències a l'hora de llegir un llibre.

3.1.1. Definició de les classes

Primer representem amb un diagrama de classes la ontologia completa, que es mostra a continuació:

Nota: no hem pogut utilitzar el plugin OntoViz a causa de que no funciona correctament a l'entorn dels laboratoris de la FIB (tal i com està especificat a la documentació de l'assignatura), per aquest motiu no hem pogut especificar la jerarquia de classes directament des del Protégé. En canvi, hem creat el diagrama de classes manualment.



A continuació farem una breu explicació de cada classe i les seves propietats.

Autor

Té tota la informació sobre els diferents autors. Els autor s'associen amb un gènere, els llibres que ha escrit i la nacionalitat que té. La informació d'aquesta classe s'obté de cada instància.

Atributs:

- **Escribe_genero:** Relació amb els gèneres en els que ha escrit algun llibre
- **NombreAutor:** El nom de l'autor
- **Ha_escrito:** Relació amb els llibres que ha escrit
- **Es_nacionalidad:** Relació amb la nacionalitat de l'autor

Genero

Conté tota la informació d'un gènere, relacionant-lo amb els autors que escriuen llibres en aquest gènere i els llibres que han estat escrits d'aquest gènere. Tota la informació d'aquesta classe s'especifica a cada instància.

Atributs:

- **Autores_genero:** Relació amb els autors que escriuen llibres d'aquest gènere
- **NombreGenero:** El nom del gènere
- **LibrosGenero:** Relació amb els llibres que formen aquest gènere

Libro

És el terme més important de la ontologia. En aquesta classe definim els llibres relacionant-los amb la temàtica del llibre, el gènere del llibre i l'autor del llibre. A les instàncies definim també informació com l'any, la dificultat, si és juvenil o no, si és bestseller o no, la valoració del llibre i la seva longitud. Tota aquesta informació ve donada per les instàncies prèviament.

Atributs:

- **Es_tematica:** Relació amb la temàtica del llibre
- **Anyo:** Any en el que es va escriure el llibre
- **Dificultad:** Valoració del 0 al 10 de la dificultat de llibre
- **Es_genero:** Relació amb el gènere del llibre
- **Juvenil:** Si és apte per a menors de 18 o no
- **Bestseller:** Si és un bestseller o no
- **Titulo:** Títol del llibre
- **Valoracion:** Valoració del 0 al 10 del llibre
- **Escrito_por:** Relació amb l'autor del llibre
- **Longitud:** Tamany del llibre en pàgines

Tematica

Defineix una temàtica i la relaciona amb els llibres d'aquesta temàtica. Ve donada per les instàncies prèviament inicialitzades.

Atributs:

- **Libros_tematica:** Relació amb els llibres d'aquesta temàtica
- **NombreTematica:** Nom de la temàtica

Nacionalidad

Conté les nacionalitats dels autors i es relaciona amb aquets. Tota la seva informació ve definida per les instàncies.

Atributs:

- **Autores_nacionalidad:** Relació amb els autors d'aquesta nacionalitat
- **NombreNacionalidad:** Nom de la nacionalitat

3.2. Mètode de resolució

A continuació farem una breu explicació dels passos seguits per a obtenir una resolució al problema.

3.2.1. Abstracció de les dades

El problema plantejat consisteix en trobar la manera de traduir una sèrie de dades proporcionades per l'usuari a fets més abstractes com poden ser les seves preferències literàries per tal de trobar les prioritats del lector.

3.2.2. Associació heurística

Les dades abstractes obtingudes anteriorment es relacionen per tal de trobar una solució abstracta. És en aquest punt en el que valorem cada paràmetre i li atorguem una puntuació favorable o no depenent del grau de relació amb les dades introduïdes per l'usuari i de si és adient o no per a la recomanació final.

3.2.3. Transformació a una solució tangible

Transformem la solució abstracta a una solució que poguem presentar a l'usuari. En el nostre cas, les recomanacions s'ordenen de manera que les que tinguin les puntuacions més altes quedin primeres. Les tres primeres són les que es presentaran a l'usuari de manera que obté les recomanacions més adients.

4. Implementació

4.1. Mòduls

- **MAIN:** Disposa de la regla inicial que inicialitza el programa, a més de contenir templates i funcions útils utilitzades pels següents mòduls.
- **recopilacion-datos-lector:** Aquest mòdul recopila informació del perfil del lector: edat, freqüència amb la que llegeix, lloc on sol llegir, si li agraden els llibres populars, si es deixa portar per les valoracions, si prefereix autors estrangers i la seva nacionalitat.

- **recopilacion-datos-preferencias:** Aquest mòdul recopila informació de les preferències del lector: gèneres preferits, temàtiques preferides i autors preferits.
- **recopilacion-datos-libros-leídos:** Aquest mòdul recopila informació dels llibres llegits pel lector, per tal d'oferir-li nous que no hagi llegit.
- **procesado-datos:** Aquest mòdul procesa tota la informació recopilada en els anteriors mòduls per escollir els llibres que formaran la nostra solució. Després de generar els facts per poder procesar la informació, puntua la col·lecció de llibres segons aquesta informació obtinguda.
- **generacion-soluciones:** Aquest mòdul construeix un llista de llibres recomanats, amb tots els llibres de la col·lecció, i els ordena primer per bestseller i després per puntuació, per tal d'oferir al lector llibres best-seller en cas que el lector no tingui preferències).
- **presentar-resultados:** Aquest mòdul conté l'última regla del programa, que presenta els resultats al lector, mostrant els tres llibres recomanats amb més puntuació.

4.2. Prototip

Al principi vam decidir realitzar la pràctica aplicant una metodologia de desenvolupament basat en prototipatge ràpid i amb disseny incremental.

L'estructura del nostre programa en CLIPS utilitzada és la següent:

- **Classes de l'ontologia:** codi produït pel programa Protégé que defineix les classes de l'ontologia utilitzada.
- **Instàncies de l'ontologia:** codi produït pel programa Protégé que defineix les instàncies de l'ontologia utilitzada.
- **Definició de classes (defclass):**
 - *Recomendacion:* per mantenir la informació dels llibres amb la seva puntuació.
 - *ListaLibrosRecomendados:* per mantenir una llista dels llibres recomanats.
- **Definició dels moduls (defmodule):**
 - *recopilacion-datos-lector:* recopila la informació del perfil del lector.
 - *recopilacion-datos-preferencias:* recopila la informació de les preferències del lector.

- *recopilacion-datos-libros-leidos*: recopila la informació dels llibres llegits.
 - *procesado-datos*: procesa la informació valorant llibres segons la informació.
 - *generacion-soluciones*: genera les solucions amb els llibres ordenats segons bestseller i valoració.
 - *presentar-resultados*: presenta els resultats al lector dels tres llibres recomanats.
- **Definició dels templates (deftemplate)**: emmagatzema la informació del perfil del lector, les seves preferències i els llibres que ha llegit, a més de la informació segons l'ordenació de la llista de recomanacions (desordenada, ordenada per best-seller i ordenada per puntuació).
 - **Definició de les funcions (deffunction)**: funcions per realitzar les preguntes al lector, a més de l'ordenació dels llibres segons bestseller o puntuació.
 - **Definició dels missatges (defmessage-handler)**: codi dels missatges utilitzats per mostrar els resultats al lector.
 - **Definició de regles (defrule)**: codi de les regles implementades per cada mòdul.

5. Jocs de prova

Tal i com ens demanàvem, hem hagut de fer un testeig del nostre programa per veure si realment els resultats que hem obtingut són coherents amb la realitat i per tant podem parlar d'un programa efectiu i ben dissenyat.

Per dur a terme aquestes proves hem delimitat un camp en el que ens centrarem, ja que en ser una practica amb una ontologia extensa creiem que no cal estudiar tots els casos així que hem decidit estudiar els següents.

- **Estudi 1**: Com influeix l'edat. Casos 1 i 2.
- **Estudi 2**: Com influeix la tria de molts o pocs gèneres. Casos 3 i 4.
- **Estudi 3**: Com influeix la quantitat d'hores que dediques a llegir a la setmana. Casos 5 i 6
- **Estudi 4**: Com influeix el coneixement dels llibres disponibles. Casos 7, 8 i 9

En cada cas d'estudi intentem només modificar els paràmetres que volem estudiar i fer que tots els altres siguin iguals i el més neutres o aleatoris possibles. Per fer això decidim que el lloc sempre serà indiferent, les lectures populars i el fet de que sigui estranger l'autor també. Les valoracions sempre les tindrem en compte i el lector serà Britànic.

A continuació us mostrem els casos plantejats, el resultat que esperem i el resultat obtingut a l'executar el programa.

Hem decidit fer un fitxer apart pels outputs de cada joc de proves perquè sigui més còmode de llegir la documentació d'aquests, incloent en aquest apartat només les conclusions extretes dels outputs.

[ESTUDI 1]

Pretenem estudiar la diferència que hi ha entre un lector menor i un major, si realment el programa ens dona unes recomanacions juvenils per al menor d'edat i no juvenils per al major d'edat.

Cas 1

Estudi del comportament d'un lector d'edat menor (edat < 18):

- **Solució esperada:** Esperem que els llibres juvenils tinguin un pes important.
- **Output:** La tria recomanada és:
 - *The Adventures of Huckleberry Finn*
 - *The Silence of the Lambs*
 - *The Testaments*
- **Comentaris de la solució:** Els llibres no tots són juvenils, per tant no és el que esperàvem.

Cas 2

Estudi del comportament d'un lector d'edat major (edat > 17):

- **Solució esperada:** Esperem que, en tria dels llibres tingui un pes important que els llibres que no siguin juvenils
- **Output:** La tria recomanada és:
 - *The Adventures of Huckleberry Finn*
 - *The Silence of the Lambs*
 - *The Testaments*
- **Comentaris de la solució:** Els llibres no tots són juvenils, per tant no és el que esperàvem.

Conclusió estudi: Els resultats no han estat els esperats ja que preveïem que l'edat seria un factor important a l'hora de triar quins llibres recomanar o no. Després de donar-li voltes

hem vist que el pes de l'edat és insignificant (comparat amb tot l'altre) i que en el cas de que totes les altres dades siguin igual ens donarà molts cops el mateix resultat. Per tant l'experiment ens ha servit per veure que com es una suma de puntuació els factors de l'edat son tant importants com qualsevol de les altres preguntes i que si les respostes son exactament iguals i només en canvia una, en aquest cas l'edat, els resultats seran iguals.

[ESTUDI 2]

Pretenem estudiar la diferència que hi ha entre triar pocs gèneres o molts gèneres, amb els altres paràmetres igual. En aquest cas i perquè no ens passi el que ens ha passat a l'estudi 1, en l'apartat d'autors preferits i llibres llegits serà aleatori.

Cas 3

Estudi del comportament d'un lector amb pocs gèneres preferits, Horror i Western (2).

- **Solució esperada:** Esperem que els llibres siguin dels gèneres escollits.
- **Output:** La tria recomanada és:
 - *The Silence of the Lambs*
 - *The Shining*
 - *The Stand*
- **Comentaris de la solució:** Com esperàvem, els llibres son d'Horror.

Cas 4

Estudi del comportament d'un lector amb molts gèneres preferits, en concret 8.

- **Solució esperada:** Esperem que els llibres siguin variats.
- **Output:** La tria recomanada és:
 - *Lord of the Flies*
 - *The Road*
 - *The Silence of the Lambs*
- **Comentaris de la solució:** Esperavem que fossin més variats ja que n'hi ha 2 que son del mateix genere.

Conclusió estudi: En aquest estudi demostrem que variant i tenint aleatorietat en certs aspectes que defineixen el coneixement del lector creem recomanacions diferents, segurament si no hagués sigut així, la tria de genres de el cas 3 i el cas 4 hi hagués conclòs amb la mateixa recomanació, però veiem que variant-los ens trobem amb casos de recomanació diferent.

El fet de que el llibre *The silence of the Lambs* hagi aparegut a la recomanació dels 4 casos que ja hem fet no sabem si és casualitat o no, en tot cas com que plantejarem més casos veurem si continua sortint i determinarem si hem de fer un estudi per veure que passa amb aquest llibre.

[ESTUDI 3]

Conclusió estudi: Aquest estudi pretén avaluar com influeix les hores que el lector dedica en el nostre programa. Com ja hem vist a l'estudi 1, no ens aporta res estudiar només un factor i per això conclourem que aquest estudi es inútil i no portarem a terme ni el cas 5 ni el 6.

[ESTUDI 4]

Pretenem estudiar el comportament del programa si no s'ha llegit cap llibre, si s'han llegit alguns, és a dir, la meitat o si s'han llegit tots.

Cas 7

Estudi del comportament d'un lector amb cap llibre llegit anteriorment.

- **Solució esperada:** Esperem que s'escullin els 3 llibres aleatoris.
- **Output:** La tria recomanada és:
 - *The Shining*
 - *The Stand*
 - *Do Androids Dream Of Electric Sheep?*
- **Comentaris de la solució:** El resultat és concloent, ja que els tres llibres són vàlids.

Cas 8

Estudi del comportament d'un lector amb la meitat de llibres llegits, del 1 al 25.

- **Solució esperada:** Esperem que s'escullin 3 llibres dels que no s'han llegit
- **Output:** La tria recomanada és:
 - *The Shining*
 - *The Stand*
 - *Do Androids Dream Of Electric Sheep?*

- **Comentaris de la solució:** El resultat és correcte ja que cap dels 3 llibres pertany al grup del 1 al 25.

Cas 9

Estudi del comportament d'un lector amb tots els llibres llegits.

- **Solució esperada:** Esperem que, com que el fet d'escollir un llibre el que fa es restar la puntuació, el resultat sigui semblant o igual al del cas 7.
- **Output:** La tria recomanada és:
 - *The Silence of the Lambs*
 - *The Shining*
 - *The Stand*
- **Comentaris de la solució:** Com esperàvem.

Conclusió estudi: En aquest estudi, a part de veure que el cas 7 i 9 coincidien perquè l'ordenació era la mateixa ja que en el cas 9 només restem 500 a cada un per tant surt igual que sense fer res; volíem veure si el programa seria compatible en temps i espai i no ens generaria cap error. Com podem veure a l'annex del joc de proves, el programa ha funcionat correctament.

Més enllà de que els resultats siguin els mateixos pels 3 casos, hem obtingut un bon estudi ja que hem vist que en cap cas ens escollirà un llibre ja llegit, excepte quan els haguem llegit tots.

[CONCLUSIÓ ESTUDIS]

Fins aquí la avaluació del nostre programa. Hem cregut que no seria proporcional provar cada cas i creiem que amb els casos fets anteriorment tenim prou certesa per demostrar que el programa és coherent i que a més funciona. Una base de dades més extensa i un millor enregistrament de les dades de cada llibre ens ajudaria a afinar més en la recomanació.

6. Conclusions

Aquesta pràctica ens ha servit per a aprendre a com conceptualitzar, formalitzar i implementar un Sistema Basat en el Coneixement que resolgui un problema que probablement solucionar-lo amb algun altre mètode seria molt costós i complexe.

Hem vist la utilitat dels sistemes experts i com crear un a partir d'un problema plantejat i utilitzant eines especialitzades com poden ser CLIPS i Protégé.

Per a la creació d'un Sistema Basat en el Coneixement, hem vist que és important anar resolent subproblemes per tal de solucionar un problema més general i seguir una metodologia clara.

Després d'un seguit de jocs de prova, veiem que el sistema resultant del nostre treball és un sistema que compleix perfectament l'objectiu, ja que a partir de la introducció de les dades per part de l'usuari, fa recomanacions adients a cada lector.

Cal dir, però, que ens hem trobat amb diversos problemes en la creació del sistema, sobretot al crear la ontologia introduint manualment totes les instàncies. Una possible millora podria ser la creació d'un codi que a través de interaccions amb APIs de webs que s'especialitzen en el llistament de llibres, s'obtinguin totes les instàncies amb una única execució del codi i s'implementin directament al programa Protégé o al llenguatge CLIPS.

Una altre millora, podria ser la de introduir més instàncies de les que ja tenim a la ontologia. Això aportaria molta més varietat a les recomanacions i més especialització. Aquesta tasca es podria fer molt més ràpidament amb el codi mencionat anteriorment.

Una última millora que podríem afegir és la introducció de més preguntes per a obtenir encara més dades sobre el lector, i, a través de proves, fer una millor parametrització de les dades obtingudes d'aquests. Això resultaria en una recomanació encara més adient per a cada lector.