

Inteligencia Artificial – Práctica 2

SISTEMAS BASADOS EN EL CONOCIMIENTO

Carlos Lázaró Costa
David Segovia Tomàs
Miriam Vall Rovira

Introducción

Esta práctica consiste en el desarrollo de un Sistema Basado en el Conocimiento, SBC, que usaremos para generar una lista de ofertas de alquileres adecuadas para varios clientes (los usuarios).

Partimos de una base de información de las distintas viviendas que están en alquiler, y según las preferencias y/o restricciones de los clientes, se determinará qué viviendas son más adecuadas para cada uno.

Seguimos las 5 fases en las que consiste la metodología explicada en clase:

- 1. Identificación del problema**
- 2. Conceptualización**
- 3. Formalización**
- 4. Implementación**
- 5. Valoración y prueba**

En los apartados siguientes se muestra el desarrollo de cada una de estas fases (siempre en función de nuestro problema) para crear el SBC. Finalmente obtenemos unas conclusiones y valoramos el trabajo hecho.

1. Identificación del problema

Sirve para determinar si el problema se puede solucionar con la creación de un SBC. Hay que comprobar que disponemos de las fuentes de conocimiento necesarias y si se puede resolver el problema dado su grado de complejidad.

1.1 Descripción del problema

El *Consorci de l'habitatge de Barcelona* quiere facilitar la vivienda más adecuada a las necesidades de las personas que la están buscando. Es un organismo que dispone de un gran número de ofertas de alquiler, cada una de las cuales tiene las características más relevantes de la vivienda. También se dispone de información sobre la ciudad de Barcelona, con sus distintos servicios (y sus coordenadas, para calcular la distancia de estos a las viviendas), y sobre los posibles clientes; las preferencias, restricciones y

características personales de cada uno de ellos.

El sistema usado para asignar viviendas está basado en esta información; y la calidad del sistema dependerá de su flexibilidad y de su capacidad para obtener las informaciones clave de los usuarios, para hacer las mejores asignaciones.

Las asignaciones de viviendas se harán en función de las características de estas, y también en función de las de los usuarios, así como las preferencias y restricciones de estos. Hay que diferenciar entre características y preferencias:

- **Preferencias del usuario:** estas dependen de cada usuario, son informaciones que se ajustan a su comportamiento real, y sirven para deducir algunas características de la vivienda más adecuada para él. También son útiles para decidir qué características de la vivienda son más importantes.
- **Características del usuario:** nos permiten clasificar a los usuarios según criterios socio-demográficos, y facilita hacerles recomendaciones. A partir de estos grupos socio-demográficos, podemos suponer preferencias o comportamientos que utilizaremos para limitar la búsqueda de viviendas.

Con los datos nombrados anteriormente se puede buscar por la información disponible para hacer las asignaciones. Por cada vivienda obtendremos diferentes características que sean recomendables para el perfil introducido de los usuarios que ocuparán la vivienda y sus necesidades aunque no lo hayan pedido explícitamente. También eliminaremos aquellas viviendas que no cumplan con las restricciones, siendo recomendaciones parcialmente adecuadas hasta que tengan 3 restricciones incumplidas lo cual hará que no se muestre esa vivienda como recomendación a no ser que no haya ninguna otra opción que cumpla más restricciones.

1.2 Viabilidad de la solución

Para decidir si es viable intentar resolver el problema desarrollando un SBC, hay que tener en cuenta qué tipo de solución se quiere ofrecer, el tamaño del dominio y la información que este contiene...

A primera vista puede parecer idóneo enfocarlo como un problema de búsqueda, porque consistiría en visitar el espacio de información escogiendo las viviendas que más se ajustan a unas restricciones y preferencias. Pero hacer esta búsqueda es demasiado costoso porque el espacio de exploración es muy grande, y si además no se dispone de conocimiento, esta exploración se realizaría de forma aleatoria.

Así que, dado que la información de la que disponemos es muy amplia, podemos resolver el problema mediante la creación de un SBC, lo cual nos permitirá obtener una solución justificada con la aplicación de reglas, restricciones y preferencias.

1.3 Fuentes de conocimiento

El SBC podrá tomar decisiones a través de conocimiento obtenido desde distintas fuentes:

En primer lugar, los **usuarios**, que además de sus características personales, indicarán cuales son sus preferencias y/o restricciones.

En segundo lugar, el **catálogo de viviendas** que se ofrecen para alquilar, el cual contiene toda la información de las viviendas disponibles, con las restricciones correspondientes.

Además, los **servicios** ofrecidos en las cercanías de las viviendas, que pueden ser de interés para los solicitantes. En función de las preferencias de éstos, serán de más o menos importancia, y también van a resultar útiles para tomar decisiones.

Finalmente, el **sentido común**, para decidir que viviendas son mejores que otras según características como la iluminación, el espacio disponible, etcétera.

1.4 Objetivos del sistema

Para poder solucionar el problema, nos tenemos que preguntar qué objetivos tiene que cumplir el SBC. Son los siguientes:

- Obtener todos los datos relevantes de los usuarios que ocuparán la vivienda, sus

restricciones y sus preferencias.

- Inferir, a parte de los servicios explícitos dados por el usuario, otros servicios ofrecidos en la ciudad a partir de los grupos de edades y estatus de los ocupantes (niños, estudiantes, jóvenes, trabajadores y jubilados), que nos será útil para valorar el grado de adecuación de cada vivienda.
- Descartar viviendas que son incompatibles para el usuario, a causa de restricciones que no se pueden cumplir o a partir de los datos inferidos de los servicios cercanos a la vivienda.
- Evaluar las posibles soluciones que no se han descartado, y determinar el grado de adecuación de cada una en función de las preferencias del usuario (que son explícitas), o de las inferidas.
- Presentar a los usuarios las viviendas más adecuadas para ellos de forma clara, con un razonamiento justificado, mostrando cómo de adecuadas son las viviendas exactamente.

1.5 Resultados esperados

El sistema generará una lista de viviendas recomendadas, según los perfiles de los usuarios. Por cada vivienda se dará el grado de recomendación, que puede ser:

- **Parcialmente adecuado:** no cumple completamente un número reducido de criterios (máximo 2), pero dependiendo del usuario se podría considerar aceptable.
- **Adecuado:** cumple todos los requerimientos.
- **Muy recomendable:** no sólo cumple con todos los requerimientos, sino que también tiene otras ventajas (por ejemplo, es más barato, más grande, tiene servicios adicionales de interés...).

Todas las soluciones adecuadas cumplen las restricciones impuestas por los usuarios.

Para priorizar sobre una recomendación u otra tenemos en cuenta las preferencias de los usuarios.

Diferenciamos entre las decisiones que se toman en base a las preferencias de los usuarios y las tomadas por información inferida de los servicios.

2. Conceptualización

En esta fase estructuramos el problema y los procedimientos para llegar a su solución.

2.1 Descripción de los elementos del dominio

El sistema debe tener esta información, separada en tres grandes grupos:

- Las viviendas, con sus características y los servicios disponibles en sus proximidades.
- Los usuarios, con sus preferencias, restricciones y características personales.
- Los servicios disponibles en la ciudad.

Características de las viviendas:

- Precio mensual
- Localización
- Superficie
- Dormitorios
- Terraza / Balcón
- Amueblada / No amueblada
- Electrodomésticos
- Tipo de vivienda (Piso / Dúplex / Unifamiliar)
- Altura (Bajo / Entresuelo / Ático)
- Altura Máxima del edificio (en caso de tipo Piso o Dúplex)
- Presencia de ascensor (en caso de tipo Piso o Dúplex)
- Prohibición de mascotas
- Calefacción / Aire acondicionado

- Piso soleado (Mañana / Tarde / Todo el día)
- Piscina comunitaria
- Plaza de garaje o parking incluida
- Vivienda con buenas vistas
- Servicios cercanos
- Otros

Características de los usuarios:

- Grupo de edad y estatus
- Tipología (con / sin pareja, con / sin hijos, familia monoparental / biparental, grupo de estudiantes, con / sin ancianos a su cargo, ...)
- Si estudia o trabaja en la ciudad
- Si posee coche
- Si tiene movilidad reducida

Preferencias y restricciones de los usuarios:

- Precio máximo que están dispuestos a pagar
- Restricciones respecto al número y tipo de dormitorios (simples o dobles)
- Precio mínimo al que tienen confianza en la oferta
- Restricciones o preferencias específicas respecto a la distancia a algún servicio
- Si necesita una vivienda adaptada a movilidad reducida
- Si requiere balcón o terraza
- Si necesita la vivienda amueblada o con electrodomésticos
- Si requiere piscina
- Si requiere calefacción o aire acondicionado
- Si tendrá mascotas en la vivienda
- Si tiene coche y necesita parking

Características de los servicios:

- Localización (Coordenadas)
- Nombre del servicio

2.2 Problemas y subproblemas

Al recopilar la información, vemos que hay que resolver, al menos, tres problemas separados. Son los siguientes:

Recopilar información de los usuarios

Hay que elaborar los datos de los usuarios, clasificándolos según datos personales o datos que indican preferencias y/o restricciones explícitas. La calidad de las recomendaciones será determinada por el grado en que se satisfacen las preferencias y restricciones.

También es necesario hacer un **análisis de las preferencias y/o restricciones**, puesto que tienen que ser evaluadas de forma razonable, completando la información extraída de estas con la inferida por las características de los usuarios.

Valoración de la información de las viviendas

Cuando tenemos guardada toda la información, disponemos de los datos de todas las viviendas y tenemos que ver si forman parte de la solución. Por lo tanto, hay que eliminar (o descartar) las viviendas que no cumplan con las restricciones de los usuarios (aunque también pueden ser descartadas si no cumplen algún requisito del sentido común). Entre las viviendas que sí cumplen las restricciones, hay que valorarlas según las preferencias de los usuarios, las explícitas y las inferidas a partir de sus características. Cuantas más preferencias cumplan, mejor será su grado de recomendabilidad, además del precio y el número de personas que pueden ocupar la vivienda.

Elaboración y presentación de una solución

Finalmente, hay que clasificar las viviendas en tres grupos: *parcialmente adecuada*, *adecuada* o *muy recomendable*. Se hará la recomendación en orden decreciente de recomendación, empezando por las *muy recomendables*. Las recomendaciones tienen que ser lo más ajustadas posibles a las preferencias de los usuarios, haciendo

distinciones entre usuarios similares.

2.3 Resolución: conocimiento extraído del dominio (deducciones)

Una vez el sistema dispone de todos los datos necesarios, es posible extraer conocimiento y principios que pueden ser aplicados en distintas ocasiones.

Se pueden clasificar las viviendas según valoraciones en diferentes aspectos, cada uno de los cuales puede tener una importancia determinada. Filtramos las viviendas respecto varios parámetros, descartando las que se rechazan explícitamente con restricciones, y ajustando las que cumplen con más preferencias (explícitas o deducidas). Así llegaremos a una solución que sea la más adecuada para el usuario.

A parte de los datos objetivos de los que disponemos, también podemos deducir información a partir de estos, como por ejemplo:

- Si se trata de una familia con hijos se valorará que estén cerca de servicios como colegios.
- Por sentido común, las viviendas se valorarán mejor si disponen de calefacción y/o aire acondicionado, y también por su cercanía a supermercados y por si son luminosas durante una buena parte del día.
- Los usuarios que tengan coche van a necesitar un lugar para aparcar, así que en ese caso se valorará la plaza de parking incluida en las viviendas.
- La gente con mascotas no pueden recibir recomendaciones de viviendas donde estén prohibidas, y si hay la posibilidad de disponer de una terraza o la vivienda está cerca de un espacio verde se va a evaluar positivamente.
- Si son mayores o tienen gente mayor a su cargo, hay que tener en cuenta la existencia de centros de salud en las proximidades de la vivienda.
- Si se trata de un estudiante o grupo de estudiantes, hay que ver qué viviendas tengan cerca bibliotecas y transporte público para ir a su lugar de estudio.
- Si trabaja se valorará el transporte público para ir a su lugar de trabajo.

2.4 Suposiciones

Se hacen algunas suposiciones:

- El número disponible de viviendas es lo bastante grande como para que, dados unos mismos datos y preferencias y/o restricciones de un cliente, se obtengan recomendaciones diferentes.
- La variedad de viviendas disponibles está equilibrada (hay viviendas suficientes de todo tipo).
- Se mantiene la prioridad de las preferencias independientemente de las características del usuario.
- Las características de las viviendas que no se consideran relevantes en un determinado caso, pueden influir en la recomendación de esas si y solo si se considera que pueden ser ventajas, en cuyo caso pasarán a afectar el grado de recomendación positivamente.

3. Formalización

En esta fase analizamos el problema para representarlo de forma adecuada implementando un SBC.

3.1 Ontología

Los conceptos del dominio que es necesario representar con clases son: Vivienda, Coordenadas (Coordinates) y Recomendación.

El esquema de relaciones entre clases es el siguiente:



3.1.2 Clase Vivienda

Es la clase que representa el concepto principal de la ontología, porque contiene la información necesaria para determinar si es una buena recomendación para un usuario. Será la información que usaremos para contrastar con las preferencias y restricciones del usuario.

Vivienda		
precio		Float
sol	Symbol	manana
		tarde
		todo-el-dia
altura		Integer
calefaccion		Boolean
electrodomesticos		Boolean
...		

Los slots de esta clase son:

- Adaptada-movilidad-reducida
- Aire-acondicionado
- Altura
- Altura-máximo
- Amueblado
- Area
- Ascensor
- Balcón
- Buenas-vistas
- Calefacción
- Dormitorios-dobles
- Dormitorios-simples
- Electrodomésticos
- Localización

- Mascotas-permitidas
- Parking
- Piscina-comunitaria
- Precio
- Sol
- Terraza
- Tipo

Se ha generado la documentación con Protegé y se ha añadido dicha documentación en formato .html dentro de la carpeta “ontología-documentación”. En esos ficheros se puede obtener una descripción detallada de cada clase y cada slot.

3.1.3 Clase CLIPS_TOP_LEVEL_SLOT_CLASS

Es una clase “falsa” (auxiliar) que contiene los mismos slots que Vivienda, y sirve esencialmente para guardar información del nivel más superior.

3.1.4 Clase Coordenadas

Esta clase sirve para representar las coordenadas que contienen información sobre los ejes X e Y de otra clase.

Coordenadas	
Y	Integer
X	Integer

3.1.5 Clase Servicio

Representa un servicio que se encuentra disponible en la ciudad, y contiene la información que hace referencia al tipo de servicio del que se trata y de su localización.

Servicio		
localizacion-servicio	Instance	Coordenadas
tipo-servicio	Symbol	centro-medico
		colegio
		supermercado
		...

Los tipos de servicio diferentes que hemos considerado han sido:

- Colegios
- Zonas verdes
- Supermercado
- Hipermercados
- Transporte público
- Bibliotecas
- Restaurante
- Centros médicos
- Discotecas
- Gimnasios

3.1.6 Dominio y cobertura (tamaño espacio de búsqueda y coste)

De entrada, podemos suponer que tenemos N viviendas y M preferencias y/o restricciones, que serán los criterios de valoración. Valoraremos las N viviendas según los M criterios, así que el coste será $O(N \times M)$. Se puede considerar bastante elevado, pero al descartar viviendas con anterioridad puede ayudar a reducir el coste de forma rápida.

El dominio de la ontología cubrirá los conceptos que nos interesa definir: Vivienda, Recomendación y Coordenadas.

Usaremos estas clases para representar las viviendas que están disponibles, los servicios que se encuentran en la ciudad y las coordenadas de estos, y las recomendaciones que hace el sistema.

3.2 Método de resolución

El mejor método para resolver el problema de hacer recomendaciones (el cual es un problema de análisis) es el de la *clasificación heurística*. En este método, escogemos una solución entre un conjunto finito de soluciones, que está formado por todas las viviendas que se ofrecen para ser alquiladas.

La clasificación heurística se divide en tres fases:

- **Abstracción de datos:** partiendo de los datos que nos proporcionan los usuarios, y habiendolos clasificado en características, preferencias o restricciones, deducimos conocimiento sobre ellos (datos abstractos).
- **Asociación heurística:** a partir de los datos abstractos que hemos deducido, pasamos a una solución abstracta. Para hacerlo, tenemos que hacer una inferencia de los datos sobre las viviendas, y una inferencia de los datos personales de las personas, no sólo infiriendo las preferencias explícitas sino también las implícitas.
- **Refinamiento (y adaptación):** en esta fase, hacemos el paso desde la solución abstracta a la concreta. Esto se hará descartando aquellas viviendas que no cumplan las restricciones explicitadas por las personas, y así garantizamos que nos vamos a mover por el espacio de soluciones válidas. Después evaluaremos las viviendas según si sus características coinciden con las preferencias de la persona (explícitas o inferidas); y si hay preferencias cumplidas, cuales (que grado de recomendación).

4. Implementación

Una vez tenemos definida la ontología, pasamos a la fase de implementación, en la que implementamos el sistema basándonos en una metodología de prototipado rápido incremental. Consiste en hacer un prototipo inicial que luego va siendo modificado

progresivamente añadiendo funcionalidades hasta llegar a uno final.

4.1 Módulos

4.1.2 MAIN

En este módulo se da inicio al sistema y, después de presentarlo, pasa al módulo de recopilación de la información.

4.1.3 Recopilación

Se recopilan los datos del usuario y sus preferencias, a través de preguntas. Es el módulo donde el sistema almacena la información que el usuario le quiera dar, respondiendo a preguntas interactivamente; por lo tanto, es en este módulo donde se hace la obtención y abstracción de la información.

4.1.4 Procesado

Aquí se hace el filtrado y procesado de la información, refinando la solución abstracta que hemos obtenido en el módulo anterior. Mediante las reglas correspondientes, se eliminan de la posible solución las viviendas que no cumplen con las restricciones.

4.1.5 Generación

Se generan los resultados, valorando las viviendas según las preferencias del usuario que se ven cumplidas, estableciendo el grado de recomendación de las viviendas. Por lo tanto, en este módulo se sigue refinando la posible solución. También se hacen justificaciones por el grado de recomendación, teniendo en cuenta las preferencias explicitadas por el usuario y las que se infieren por sus características.

4.1.6 Presentación

Es el módulo que hace la presentación de las soluciones, que serán las viviendas con mayor grado de recomendación para los usuarios, junto con la justificación y la lista de posibles criterios que la vivienda no cumple, en caso de que los haya.

4.2 Prototipos

Cuando ya se había hecho una implementación de la ontología, se diseñó un prototipo inicial, al que se le añadieron funcionalidades, y también se crearon unas cuantas instancias bastante sencillas para ir haciendo pruebas. Las primeras funcionalidades fueron a partir del precio, seguidas del número de dormitorios y finalmente se añadieron las inferencias, otras recomendaciones de características de la vivienda y por último los servicios cercanos.

4.2.1 Prototipo final

Después de las pruebas y de ir ampliando sucesivamente la ontología, se llegó al prototipo final, que tiene la estructura siguiente (en ese orden):

- Clases
- Instancias
- Módulos
- Templates
- Mensajes
- Funciones
- Reglas

5. Juegos de pruebas

Verificaremos el funcionamiento del SBC. Pensamos en entradas para el sistema que representan unos cuantos casos bastante comunes y otros que no lo son tanto.

En todos los casos, aunque varían las preferencias y restricciones de cada usuario, se considerarán prioritarias aquellas viviendas que cumplan con condiciones que, según el sentido común, las hacen mejores que el resto. Por ejemplo, tener supermercados, gimnasios o restaurantes cerca, o que la vivienda sea iluminada por el sol durante buena parte del día.

5.1 Primer caso básico: Familia con hijos

Para una familia con hijos, esperamos unos resultados determinados. Por un lado, es lógico esperar que se les recomienden viviendas unifamiliares o pisos muy espaciosos, además de más de dos habitaciones (y por lo menos una doble). También se espera que la vivienda esté cerca de zonas verdes, en un lugar tranquilo y lejos de zonas de ocio nocturnas. Además, sería necesaria la presencia de como mínimo un hipermercado en las proximidades. Finalmente, cabe esperar que la familia posea un coche, por lo que seguramente será necesario disponer de una plaza para aparcar.

Juego de prueba se encuentra en el fichero .txt

5.2 Segundo caso básico: Estudiante

En este caso, ya que por lo general no se suelen tener ingresos muy altos, se espera que el precio máximo que está dispuesto a pagar sea bajo, y que la vivienda no tenga muchos metros cuadrados, porque no son necesarios. Probablemente le interesará estar cerca de zonas nocturnas de ocio.

Juego de prueba se encuentra en el fichero .txt

5.3 Tercer caso básico: Matrimonio de edad avanzada

Si se trata de personas con edad avanzada, es muy probable que sea conveniente la presencia de centros de salud cerca de la vivienda, y también que esta esté situada en una zona tranquila.

----- Sistema de Recomendación de Viviendas -----

Escoge el precio mínimo: [Mínimo 1] 100
Escoge el precio máximo: [Mínimo 100] 1000
¿Estarías dispuesto a pagar más si la oferta lo merece? [si][no]: si
¿Te gustaría indicar un mínimo de metros cuadrados? [si][no]: si
¿Cuántos metros cuadrados quieres como mínimo? [Mínimo 1] 100
¿Cuántas personas habrá en casa? [Mínimo 1] 2
¿Cuántos dormitorios dobles necesitas? [Mínimo 0] 1
¿Quiere terraza obligatoriamente? [si][no]: no
¿Quiere balcón obligatoriamente? [si][no]: no
¿Necesita que la vivienda esté amueblada? [si][no]: si
¿Necesita la vivienda con electrodomesticos? [si][no]: si
¿Necesita que la vivienda tenga calefacción? [si][no]: si
¿Necesita la vivienda con aire acondicionado? [si][no]: si
¿Quiere que la vivienda tenga piscina comunitaria o propia? [si][no]: no
¿Alguno de los ocupantes tiene movilidad reducida? [si][no]: si
¿Tiene mascotas o tiene pensado tenerlas en el futuro? [si][no]: no
¿Le gustaría disponer de parking en el mismo edificio de su vivienda? [si][no]: no
¿Necesita algún servicio cercano a la vivienda? [si][no]: si
¿Requiere que alguno de los siguientes servicios esté cerca de la vivienda?
1. centro-medico
2. colegio
3. supermercado
4. hipermercado
5. zona-verde
6. discoteca
7. transporte-publico
8. gimnasio
9. biblioteca
10. restaurante
Indica los números separados por un espacio: 7
¿Habrá niños en la vivienda? [si][no]: no
¿Alguno de los ocupantes trabaja? [si][no]: no
¿Alguno de los ocupantes estudia? [si][no]: no
¿Hay algun ocupante entre 20 y a 30 años? [si][no]: no
¿Hay algun ocupante mayor a 60 años? [si][no]: si

La salida está en el .txt

6. Conclusiones

En conclusión, en esta práctica hemos creado un sistema basado en el conocimiento que hace recomendaciones a usuarios según las preferencias y/o restricciones de estos. Estas recomendaciones se hacen sin tener en cuenta posibles recomendaciones que se hayan podido hacer con anterioridad. El SBC ha sido desarrollado siguiendo la metodología vista en clase.

Hemos hecho pruebas con casos básicos y otros más extremos, y para ambos tipos se ha comportado como esperábamos que lo hiciera, así que podemos afirmar que cumple con

los objetivos.

Respecto a las posibles mejoras del sistema, es evidente que como más amplia fuese la ontología, más completo sería el sistema, porque se harían las recomendaciones en función de muchos parámetros muy variados, y serían más ajustadas (y por lo tanto, mejores). Una forma de ampliar la ontología sería añadirle slots a la clase Vivienda, aumentando las posibles características a tener en cuenta de una vivienda. Otra manera sería hacer más preguntas al usuario para recopilar más información sobre él y obtener mejores deducciones. También se mejoraría añadiendo nuevos servicios. En todo caso, cualquier ampliación de la ontología se traduciría a un aumento de reglas.