

SEBREI: Sistema Experto basado en reglas para la recomendación de equipos de impesión *

Erica Vidal

Fac. de Cs. Exactas, Ingeniería y Agrimensura
Departamento de Ciencias de la computación
Pellegrini 250, Rosario, Santa Fe, Argentina
ericavidal@gmail.com

1. SISTEMAS EXPERTOS

Los términos sistema experto, sistema basado en conocimiento o sistema experto basado en conocimiento a menudo se usan como sinónimos.

Los sistemas expertos son una rama de la Inteligencia Artificial que hace un amplio uso del conocimiento especializado para resolver problemas como un especialista humano.[7]

Se los puede definir como una clase de programas que son capaces de: aconsejar, categorizar, analizar, consultar, diseñar, diagnosticar, explicar, explorar, interpretar, justificar, planificar, son en suma, programas capaces de manejar problemas que normalmente requieren para su resolución la intervención humana especializada.[6]

Un experto es una persona que tiene experiencia en cierta área, tiene habilidades que la mayoría no y/o puede resolver problemas que la gran parte de las personas no pueden, o lo puede hacer de una manera más eficiente.

El conocimiento en un sistema experto puede representar experiencia o conocimiento que está disponible en libros, revistas y personas.

Todo sistema experto está especializado en cierto campo o área del conocimiento. El área de conocimiento que es capturado por un sistema experto se llama dominio de tarea o aplicación.

Estructura básica de un Sistema Experto

La estructura básica de un sistema experto consta de una base de conocimiento que es el conjunto de hechos y heurísticas debidamente almacenadas y un motor de inferencias que es el programa que gestiona y controla la base de conocimiento. Una característica relevante es que la base de conocimiento y el motor de inferencias se encuentran separados. En la figura 1 se ilustra la estructura básica de un

sistema experto.

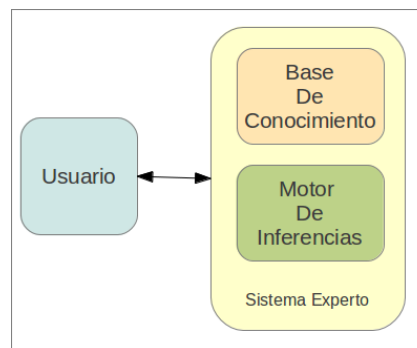


Figure 1: Estructura básica de un sistema experto.

Otros componentes importantes son la memoria de trabajo, la interfaz con el usuario y el módulo de explicación. La memoria de trabajo también llamada base o lista de hechos (facts list) es el conjunto de datos de un determinado problema y se almacena en forma separada de la base de conocimiento. El módulo de explicación o justificación explica el razonamiento utilizado por el sistema para llegar a una conclusión.

Representación del conocimiento

“En IA, la representación del conocimiento es una combinación de estructura de datos y procedimientos de interpretación, que si se usan de una forma adecuada en un programa se tendrá un comportamiento experto.”[5]

La representación del conocimiento formaliza y ordena el conocimiento. Existen varios sistemas para la representación del conocimiento. Entre los principales sistemas se tienen:

- Formalismos lógicos (lógica proposicional, lógica de primer orden, lógicas multivaluadas)
- Sistemas de producción
- Formalismos estructurados (Redes semánticas, Frames, Objetos y Ontologías)

Reglas de Producción

Una regla de producción es una sentencia que tiene la siguiente forma:

SI <antecedente> ENTONCES <consecuente>

donde el antecedente y el consecuente son frases en lenguaje coloquial. El nombre de producción se debe al hecho de que el consecuente describe el estado de cosas que resulta como consecuencia de que el antecedente sea satisfecho.

Un sistema basado en reglas se define como una colección consistente de reglas del tipo antecedente-consecuente, una base de hechos y un motor de inferencias. [9]

Motor de inferencia

El motor de inferencia de un sistema experto usa los datos (hechos o evidencia) y el conocimiento (el conjunto de reglas almacenado en la base de conocimiento) para obtener nuevas conclusiones o hechos mediante el uso de reglas de inferencia.

Una regla de inferencia es un esquema para construir inferencias válidas. Estos esquemas establecen relaciones sintácticas entre un conjunto de fórmulas llamados premisas y una aserción llamada conclusión. Ejemplos de reglas de inferencias clásicas son las conocidas Modus Ponens y Modus Tollens.

Agenda

La lista de reglas que se encuentran en la base de conocimiento que son potencialmente ejecutables se almacenan en lo que se denomina agenda. Una regla es potencialmente ejecutable si su parte izquierda se satisface. Al conjunto de reglas que pueden ser ejecutadas se las denomina grupo de conflicto de reglas y al preproceso de ordenarlas para ser ejecutadas se lo denomina resolución de conflictos.

El algoritmo de Rete

El motor de inferencias debe decidir que regla disparar según el contexto, el componente del motor de inferencia que matchea (compara patrones) hechos con patrones en reglas para determinar que reglas tienen sus condiciones satisfechas deberá buscar entre posiblemente miles de reglas. Existen un conjunto de algoritmos que abordan esta problemática de manera muy eficiente, el más conocido es el algoritmo de Rete.

Los lenguajes basados en reglas tienen lo que se llama redundancia temporal, que significa que en cada ciclo de ejecución sólo pocos hechos se agregan a la lista de hechos y los mismos afectan a pocas reglas; por lo que el sistema cambia muy lentamente. El algoritmo de Rete toma ventaja de esto, salvando el estado del matching ciclo a ciclo y reconstruyendo los cambios en el estado solo para los cambios que ocurrieron en la fact list.

Encadenamientos

Un grupo de múltiples inferencias que conecta un problema con una solución se llama encadenamiento (chain).

Una encadenamiento que va desde los hechos hacia un objetivo o conclusión se llama encadenamiento hacia adelante o Forward Chain, mientras que un encadenamiento que va desde una hipótesis u objetivo hacia los hechos que soportan

dicha hipótesis se llama encadenamiento hacia atrás. Backward Chain.

Debido a que los datos determinan que reglas se usarán el Forward Chaining es llamado inferencia data-driven en contraste con inferencia goal-driven referido al Backward Chaining.

Agregar ejemplos

Suele resultar útil visualizar los Forward y Backward Chaining en términos de la búsqueda de un path a través de un espacio de búsqueda.

Buenas aplicaciones para Forward Chaining son aquellas en el que el árbol de búsqueda es ancho y no muy profundo, porque el Forward Chaining produce una búsqueda a lo ancho.

En cambio el Backward Chaining produce una búsqueda en profundidad y un buen árbol para una búsqueda en profundidad es uno angosto y profundo.[7]

De acuerdo al tipo de encadenamiento que use un motor de inferencia se los denomina también Forward o Backward Chaining.

El algoritmo de Forward Chaining trabaja de la siguiente manera:

procesa un hecho en la memoria de trabajo, verifica su base de reglas para ver si alguna de las condiciones es verdadera. Si alguna lo es, las acciones del lado derecho de las reglas son procesadas, pudiendo cambiar la información de la memoria de trabajo. Cuando se completa la acción, el motor chequea si la condición de alguna otra regla es verdadera y el proceso se repite hasta que no existan reglas que sean verdaderas.

!!algoritmo

El algoritmo de Backward Chaining en cambio:

Se comienza con una conclusión que el motor trata de satisfacer. Si no se puede satisfacer, se buscan conclusiones intermedias que sí se puedan satisfacer (sub goals), lo que permite satisfacer alguna parte del objetivo actual (current goal). Se continúa con este proceso hasta que se prueban las conclusiones iniciales o no hay más sub-conclusiones (sub goals) para analizar. [8]

!!algoritmo

Metodología para el desarrollo de sistemas expertos

“Desde los primeros Sistemas Expertos operativos desarrollados en la Universidad de Standford en la década de los ’60 hasta los que hoy se desarrollan comercialmente, su proceso de construcción ha evolucionado desde lo artesanal en aquellos días hasta un sólido enfoque metodológico en la actualidad.”[6]

A través de la bibliografía consultada sobre la construcción

de sistemas expertos, se ha observado que al igual que los modelos tradicionales de desarrollo de software, que se han propuesto para controlar el ciclo de vida del software, (como el modelo de cascada, el modelo incremental, el modelo de espiral, etc); también se han desarrollado varias metodologías para la producción de sistemas expertos.[7],[3]

El ciclo de vida del software o “los modelos de procesos son en realidad metametodologías porque determinan el orden y la duración en el cual los métodos de software comunes son aplicados. Los métodos tradicionales de desarrollo de software muestran procedimientos específicos para llevar a cabo una las siguientes etapas”[7]¹.

- Ingeniería de requerimientos
- Diseño
- Implementación
- Evaluación (verificación + validación)
- Documentación
- Mantenimiento

Además de estas, el desarrollo de un sistema experto debe incluir las etapas de

- Adquisición y representación del conocimiento
- Selección de las herramientas de desarrollo

Adquisición del conocimiento

Uno de los aspectos más importantes en el proceso de construcción de un sistema informático basado en el conocimiento es el de la adquisición y validación de conocimiento.

”Dentro de los métodos de adquisición de conocimiento se pueden citar los métodos basados en interacción humana tales como tareas familiares, entrevistas, tareas de proceso restringido y tareas de información limitada y los basados en técnicas de aprendizaje automático.”[6]

En [3] el autor en el desarrollo de su metodología integral para el desarrollo de sistemas basados en conocimiento, explora, recopila y analiza varias técnicas para la adquisición del conocimiento y la elección adecuada de su representación.

En reglas generales la adquisición del conocimiento consta de identificar las fuentes del conocimiento, extraer el conocimiento de dichas las fuentes, ordenarlo, construir definiciones claves del dominio de aplicación.

Herramientas

Las herramientas informáticas utilizadas para la realización de sistemas expertos pueden ser lenguajes de alto nivel (C, JAVA) o ambientes especializados para aplicaciones de inteligencia artificial (Common LISP, PROLOG) o ambientes

especializados para la creaciones de sistemas expertos que en la literatura se suelen denominar shells.[2]

Un shell es un programa preparado para la generación de sistemas expertos, incorporan un motor de inferencia y el ingeniero del conocimiento se limita a crear la base de conocimiento. CLIPS es un ejemplo de shell o entorno de desarrollo para la producción de sistemas expertos. En [1] hay una extensa referencia a los distintas herramientas y shell disponibles hacia la década de los '90.

LISP: (LISt ProceSIng) Inicialmente instrumentado de 1958 a 1960 por Mc Crathy, basandose en la lógica matemática. Hacia mediados de la década del 80 el 95% de los programas en IA estaban escritos en LIPS. [4]²

Common LISP :Durante los años 1980 y 1990, fue hecho un gran esfuerzo para unificar los numerosos dialectos del Lisp en un solo lenguaje (más notablemente, InterLisp, MacLisp, ZetaLisp, MetaLisp, y Franz Lisp). El nuevo lenguaje, Common Lisp, fue esencialmente un subconjunto compatible de los dialectos que reemplazó. En 1994, la ANSI publicó el estándar del Common Lisp, “ANSI X3.226-1994 Information Technology Programming Language Common Lisp Common Lisp es un lenguaje multi paradigma de propósitos generales. Soporta una combinación de paradigmas de programación como procedimental (imperativo), funcional, y orientada al objeto.

PROLOG (PROgrammation en LOGique) es un lenguaje de programación lógico e interpretado.

CLIPS (C Language Integrated Production System) es una herramienta para el desarrollo de sistemas expertos (SE) creada por la Software Technology Branch (STB), NASA / Lyndon B. Johnson Space Center. Los orígenes de CLIPS se remontan a 1984. En la actualidad, entre los paradigmas de programación que soporta CLIPS se encuentran la Programación lógica, la Programación imperativa y la Programación Orientada a Objetos. Aunque ahora es de dominio público (license public domain), aún es actualizado y mantenido por su autor original, Gary Riley. Permite representar reglas de producción y frames. El motor de inferencias es de tipo Forward Chaining, y hay X formas de resolución de conflicto (depth, breadth, LEX, MEA, complexity, simplicity y random). Está escrito en C, por lo que es portable a varios sistemas operativos. En la actualidad se encuentra en su versión 6.3.

Existen derivados de CLIPS para ser usados en otros lenguajes como por ejemplo:

- JESS: implementación de CLIPS en Java+.
- FuzzyCLIPS: incorpora a CLIPS la posibilidad de usar razonamiento difuso.
- CLIPSM: una interfaz libre de CLIPS con C++.
- PHLIPS: extensión para PHP. (PHP License V3)
- PyCLIPS: interfaz libre para Python., (utilizada en este trabajo).(License (LGPL))

¹pag. 318-319

²pag. 33

- CLIPSNET: librería de C#

Hay una extensa lista de softwares y variates derivadas de Clips y aplicaciones desarrolladas con Clips en el sitio web de la herramienta³.

SOAR (State, Operator And Result) Soar es una arquitectura general de desarrollo de agentes inteligentes. Es usada tanto por investigadores de inteligencia artificial como de ciencias cognitivas para una variedad de tareas. SOAR se usa desde 1983 y en la actualidad a evolucionado hasta su versión actual 9.3. Se distribuye bajo licencia BSD.

References

- [1] Bourbakis, N.G. *Knowledge Engineering Shells: Systems and Techniques*. Advanced Series on Artificial Intelligence. World Scientific, 1993.
- [2] Cammarata, Silvio. *Sistemi Esperti: teorie, metodi, strumenti tecnici*. Milano, ETAS Libri, 1987.
- [3] Carrillo Verdún, José Domingo. *Metodología para el desarrollo de sistemas expertos*. PhD thesis, 1987.
- [4] Charniak, Eugene; McDermott, Drew. *Introduction to Artificial Intelligence*. Addison-Wesley, 1985.
- [5] Feigenbaum, Edward A. ; Barr, Avron. *The Handbook of Artificial Intelligence*. Stanford, California, HeurisTech Press; Los Altos, California, W. Kaufmann Inc., 1982.
- [6] Garcia Martinez, Ramon; Britos, Paola Veronica. *Ingeniería de Sistemas Expertos*. Nueva Librería, 2004.
- [7] Giarratano, Joseph; Riley, Gary. *Sistemas Expertos, Principios y Programación*. PWS Publishing Company, third edition, 1998.
- [8] Silveira, Luciano. *Diseño en implementación de un motor de reglas dinámicos usando especificaciones GeneXus: Tesis de Maestría* UNIVERSIDAD DE LA REPUBLICA, Uruguay - Montevideo, 2010.
- [9] Waterman, D. A. *User-Oriented System for capturing Expertise: A rule based Approach*. Edingburg University Press, 1984.

DESARROLLO DEL SISTEMA EXPERTO

Desarrollo del sistema experto

1. PRESENTACIÓN DEL PROBLEMA

La elección de una determinada tecnología de impresión a la hora de comprar está sujeta a numerosas cuestiones relacionadas con el uso que se le va a dar al equipo, la velocidad de respuesta deseada, la calidad esperada, factores económicos, etc. Requiere tiempo y esfuerzo investigar las tecnologías disponibles en el mercado y decidir en base a las características que posee cada equipo, aquel que resulte más conveniente según las necesidades.

Los sitios web de las diferentes marcas de impresión a menudo preclasifican sus productos para ayudar al cliente a encontrar el equipo que se necesita, otros simplemente listan los productos disponibles. En todos los casos, el cliente debe saber qué está buscando. Para facilitar la tarea de elección, lo más común es consultar con un vendedor competente en la materia.

Emular un vendedor o recomendador de impresoras es un problema bastante complejo si se lo quisiera desarrollar como una aplicación convencional, además cada experto tiene su opinión sobre cuál es la mejor opción en cada caso para satisfacer al cliente.

1.1 Objetivos

El principal objetivo del presente trabajo es diseñar y desarrollar un Sistema Experto online que emule un vendedor de sistemas de impresión, que aconseja a un cliente sobre las características necesarias en impresora o multifunción de acuerdo a las necesidades que el cliente exprese tener, y en base a las características recomendadas, ofrezca de un conjunto de productos disponibles aquellos que más se ajusten.

Un segundo objetivo es la confección de un paquete de software de código abierto con la implementación del sistema.

1.2 Requerimientos generales

1.2.1 Uso

El sistema debe poder ser accedido a través de un browser. El sistema debe recomendar los equipos de impresión que más se ajusten a las necesidades del cliente. Para determinar las necesidades del cliente el sistema se vale de preguntas en español. El sistema debe explicar las razones de sus recomendaciones.

1.2.2 Actualización y Mantenimiento

El sistema debe permitir a un usuario autenticado agregar nuevos equipos o modificar los ya existentes. También se debe permitir agregar o modificar las características técnicas posibles de los equipos de impresión.

1.2.3 Cambios futuros

Se debe prever la posibilidad de extender la base de conocimiento (hechos y reglas) sin afectar la estructura del sistema.

2. ADQUISICIÓN DEL CONOCIMIENTO

2.1 Fuentes de conocimiento

³<http://clipsrules.sourceforge.net/OtherWeb.html>

La adquisición del conocimiento se llevó a cabo consultando las especificaciones técnicas publicadas por los fabricantes de equipos de impresión y a través de, las interacciones con un experto en el rubro de venta de equipos de impresión para oficina y el hogar, la página web de la empresa del experto⁴ y la web de varias marcas de impresoras y multifunción^{5, 6, 7, 8}. También se consultaron foros, noticias de tecnología y recomendaciones online publicadas por expertos en medios digitales[1],[2],[3], [4].

2.2 Definiciones y conceptos clave del dominio de tarea

2.2.1 Impresora - Multifunción:

Una impresora es un periférico de ordenador que permite producir una gama permanente de textos o gráficos de documentos almacenados en formato electrónico, imprimiéndolos en medios físicos, normalmente en papel o transparencias, utilizando cartuchos de tinta o tecnología láser. Si además de imprimir el equipo cuenta con un escáner y/o fax incorporados, el equipo se denomina multifunción.

2.2.2 Funciones:

Una impresora posee la función de imprimir, pero si el equipo es una multifunción significa que además posee una o más de las siguientes funciones: escanear, copiar, envío de fax, envío digital. El envío digital es la capacidad enviar por mail un documento digital que el equipo tenga en memoria, por lo general como resultado de haber escaneado algo.

2.2.3 Tecnologías de impresión:

Matriz de punto. Las impresoras de matriz de puntos usan un conjunto de agujas de cabeza redonda que presionan una cinta entintada contra una página. Las agujas (por lo general 9 o 24) están dispuestas en una cuadrícula rectangular llamada matriz; diferentes combinaciones de agujas forman los distintos caracteres e imágenes. Siguen siendo las únicas capaces de imprimir en formulario continuo. No procesan los documentos una página a la vez. En su lugar, trabajan principalmente con un flujo de caracteres ASCII hasta una línea a la vez y, por lo tanto, requieren de búferes de memoria muy reducidos. Como resultado, su velocidad se mide en caracteres por segundo (cps) en vez de en páginas por minuto.

Inyección de tinta. Las impresoras de inyección de tinta tienen boquillas diminutas situadas en lo que se llama cabezal de impresión, que esparcen tinta especialmente formulada sobre una página. Existen dos métodos de impresión con inyección de tinta, uno es térmico como la que usa la línea Bubblejet de Canon y HP, en los que diminutos elementos calentadores son usados para expulsar gotitas de tinta desde las boquillas del cabezal. El otro método utiliza la tecnología piezoeléctrica desarrollada por Epson.

Láser: Las impresoras láser funcionan creando una imagen electrostática de una página completa sobre un tambor fotosensible con un haz de luz láser. Cuando se aplica al tambor

el polvo ultrafino de color, denominado tóner, éste se adhiere sólo a las áreas sensibilizadas correspondientes a las letras o imágenes sobre la página. El tambor gira y se presiona contra una hoja de papel, transfiriendo el tóner a la página y creando la imagen. Esta tecnología es similar a la que utilizan las fotocopadoras. Muchas de las llamadas impresoras láser son actualmente del tipo LED. Estas impresoras reemplazan el haz de luz láser con una disposición fija de diodos emisores de luz (LEDs) para crear la imagen; por lo demás, son similares en desempeño aunque más económicas.

Tinta Sólida: Las impresoras de tinta sólida son impresoras de página completa que usan varillas de tinta encerada sólida en un proceso "phase-change" (cambio de fase). Trabajan licuando las varillas en depósitos, y luego volcando la tinta en un tambor de transferencia, desde donde es fusionada en frío en el papel en una sola pasada. Es una tecnología exclusiva de Xerox.

Sublimación de tinta: Las impresoras de sublimación de tinta son dispositivos especializados ampliamente usados en aplicaciones fotográficas y de artes gráficas. Estas impresoras trabajan calentando la tinta hasta convertirla en gas. El elemento térmico puede generar diferentes temperaturas, lo que permite controlar la cantidad de tinta que es ubicada en una mancha. En la práctica, esto significa que el color es aplicado como un tono continuo más que como puntos. Un color por vez es depositado en toda la hoja, comenzando con el amarillo y terminando con el negro. La sublimación de tinta requiere de un papel especial particularmente caro, y los tintes están diseñados para difuminarse en la superficie del papel, mezclándose para crear sombras de colores precisas.

Móvil: Las impresoras portátiles suministran muchos de los recursos que ofrecen sus similares, pero son lo suficientemente ligeras como para llevarlas a cualquier parte. Por lo general incluyen una batería recargable, y pueden ser térmicas o a inyección de tinta.

2.2.4 Velocidad de impresión:

La velocidad de una impresora suele medirse con el parámetro *ppm* (páginas por minuto), es decir, la cantidad de páginas que puede imprimir en un minuto en determinadas condiciones. Otro parámetro que se utiliza es el de *cps* (caracteres por segundo) adecuado para las impresoras matriciales que aún se fabrican. La velocidad exacta varía según la configuración del sistema, la aplicación de software, el driver y la complejidad del documento. Para poder comparar impresoras de distintos fabricantes basándose en las ppm se definió el estándar ISO/IEC 24734:2009 [5] aplicables tanto a impresoras como multifunciones que pueden imprimir en formato A4.

2.2.5 Calidad de impresión:

Uno de los determinantes de la calidad de la impresión realizada, es la resolución que se usa para describir la agudeza y claridad de la salida impresa. La resolución de impresión se mide en dpi (dots per inch) o en español, ppp (puntos por pulgada). Una resolución de "300 dpi" se refiere a que en cada pulgada (2.54 cm) cuadrada, la impresora puede situar 300 puntos horizontales y 300 verticales. Una expresión del tipo "600 x 300 dpi", significa que el primer valor se asume

⁴<http://www.jeanmco.com.ar>

⁵<http://www8.hp.com>

⁶<http://global.latin.epson.com/>

⁷<http://www.xerox.com>

⁸<http://www.brother.es/>

a la línea horizontal y el segundo a la vertical. Otro determinante de la calidad de impresión es el número de niveles o graduaciones que pueden ser impresos por punto, una técnica de capas de color que hace que la oscilación en los gráficos y fotografías sea más difícil de ver. Las impresoras sin niveles de impresión por punto, imprimen cada punto de color en una de sólo dos intensidades (encendido o apagado), con tinta cian, magenta, amarilla o negra. Pueden combinarlas para crear tintas roja, verde y morada, y pueden crear la ilusión de otros colores al distribuir puntos de distintos colores en el papel (cada color se logra siguiendo un patrón determinado). La impresión de multinivel hace posibles más intensidades para cada punto que se imprime, así permite que la impresora utilice menos puntos para crear colores esfumados y hace que sea más difícil ver los patrones.

Máquinas con la misma resolución pueden ofrecer resultados dispares, porque hay que tener en cuenta el tamaño de las gotas que generarán esos puntos por pulgada y ésta varía según la tecnología empleada para llevar a cabo la impresión. Las gotas de tinta se miden en picolitros (1 picolitro es la billonésima parte de un litro 10^{-12}).

2.2.6 Memoria:

La mayoría de las impresoras contienen una cantidad de memoria para almacenar parte de la información que se les va enviando, de forma que la computadora no tiene que estar esperando a la impresora.

2.2.7 Conectividad / Interfaces:

La conectividad o conexión refiere a la forma en que la impresora se comunica con el medio, mientras que las interfaces son los conectores que permiten la existencia de una conexión. La forma más antigua de conexión era mediante puerto serie en donde la transferencia se hacía bit a bit, permitía distancias largas con velocidades lentas que no superaban los 19.200 bytes/segundo. Actualmente una de las formas de conexión puede establecerse mediante puerto paralelo en la que las transferencias eran byte a byte permitiendo 8 conexiones paralelas consiguiendo una velocidad más rápida entre los 0.5 MB/segundo hasta los 4MB/segundo. La longitud del cable que une la impresora con el computador o print server no puede ser mayor a 2 metros.

Otra método de conexión más actual es por medio de puertos USB (Universal Serial Bus).

A algunas impresoras se las puede conectar poniéndolas en red Ethernet mediante conexiones RJ 45 basadas en el estándar IEEE 802.3 o vía wifi, mediante el protocolo IEEE 802.11. También como forma de conexión sin cables es a través de Bluetooth.

Los equipos que pueden enviar fax cuentan con una interfaz RJ 11 para poder ser conectados a la línea telefónica.

Existe otro tipo de conexión que se denomina directa y está relacionado con las interfaces para conectar pendrives, tarjetas de memoria o cámaras fotográficas sin la intervención de la computadora.

2.2.8 Lenguajes de descripción de páginas:

Hace unos años el texto a imprimir en un impresora era enviado en código ASCII con un caracter indicando bold, itálica, tamaño, etc; y los gráficos eran producidos línea por línea. El problema es que este proceso era muy dependiente del dispositivo y los caracteres raramente estaban en la misma posición con que, aparecían en pantalla. Actualmente se usa un lenguaje de descripción de página (PDL), el cual describe el formato de las páginas, la colocación del texto y los elementos gráficos como mapas de bits o como objetos vectoriales.

Existen muy variados PDLs⁹, los más conocidos son PCL¹⁰ de Hewlett Packard y Postscript de Adobe.

Postscript fue uno de los primeros PDL estandar y multi-plataforma; describe las páginas en forma vectorial en outline, las cuales son enviadas a la impresora para ser rasterizadas¹¹ en el mismo dispositivo.

PCL es un lenguaje de descripción de página desarrollado por Hewlett-Packard, a principios de los ochenta, para sus impresoras. El PCL se ha convertido en un estándar emulado por muchos otros fabricantes de impresoras. Además del texto real que se imprime, el PCL consta en gran medida de comandos diseñados para accionar diversas características y capacidades de la impresora.

Lenguajes GDI: Algunas impresoras de inyección de tinta o de matriz de puntos no usan ninguno de los PDL "clásicos" (PostScript o HP-PCL), sino que recurren a la computadora para producir la página a imprimir. A estas impresoras se les llama impresoras basadas en el host o servidor de impresora. Algunas variantes de este tipo de impresión incluyen impresoras que usan el motor de la interfaz de dispositivo gráfico (GDI) de Windows para generar la imagen de la página (impresoras GDI) y la línea de impresoras PPA (Arquitectura para Rendimiento de Impresión) de Hewlett-Packard."[3] Esta característica es importante porque dependiendo de cómo esté diseñada la impresora basada en el host, podría no ser posible imprimir desde sistemas operativos distintos a Windows.

Emulaciones: Decimos que una impresora "emula" cuando admite códigos y modos de funcionamiento de otros modelos o marcas del mercado.

2.2.9 Ciclo mensual de trabajo - Volumen de impresiones mensuales recomendados:

Los fabricantes de equipos de impresión, normalmente indican el ciclo de trabajo mensual, que es la cantidad máxima de páginas en un determinado formato que el equipo puede imprimir en un mes sin deteriorarse. Lo recomendado es que al momento de elegir un equipo el cliente elija uno cuyo ciclo de trabajo mensual exceda por mucho a la cantidad de impresiones que se van a hacer por mes. Por eso es que paralelamente, los fabricantes, también indican un volumen de páginas mensuales recomendado que es el que debe usarse para determinar cuál es el tamaño que se necesita.

⁹http://www.undocprint.org/formats/page_description_languages

¹⁰Printer Command Language.

¹¹Convertidas en puntos.

2.2.10 Color

.. Algunas impresoras utilizan dos cartuchos, uno para la tinta negra y otro para la de color, en donde suelen estar los tres colores básicos.

2.2.11 Consumibles

¹² En las impresoras matriciales, el mantenimiento es sencillo y normalmente barato, ya que únicamente será preciso cambiar la cinta, que suele durar bastante. El cabezal también se suele considerar como consumible, pero no es algo que se estropee con frecuencia, teniendo una media de vida del orden de años.

En las de inyección, lo que más cambiaremos serán los cartuchos de negro y/o color; el cabezal inyector no suele estropearse, y la mayoría de impresoras implementan sistemas de limpieza y calibrado que permiten ajustar el funcionamiento de las mismas.

Por último, en las impresoras láser el consumible por excelencia es el cartucho de tóner (es decir, la tinta), pero tampoco es extraño tener que cambiar la unidad fotoconductora o el tambor (en algunas máquinas estas dos piezas se engloban en una sola) cada cierto tiempo, siempre según la carga de trabajo que soporte la máquina.

2.2.12 Rendimiento - Consumibles

El rendimiento por página indica la cantidad total de páginas que pueden imprimirse con un cartucho. En diciembre de 2006, la Organización Internacional de Estandarización (ISO, por sus siglas en inglés) aprobó nuevos estándares para la medición del rendimiento por página de cartuchos de tinta (ISO/IEC 24711) y tóner de color (ISO/IEC 19798) para impresoras. Estos estándares complementan la norma de rendimiento ISO/IEC 19752 para cartuchos de tóner monocromático publicada en junio de 2004. ¹³

Estos estándares permiten comparar el rendimiento de cartuchos de distintas marcas.

2.2.13 Grandes formatos

Un plotter es un periférico de salida que efectúa con gran precisión, impresiones gráficas que una impresora no podría obtener. Permite imprimir mapas de gran formato, dibujos técnicos, pósters, fotografías en materiales como vinilo adhesivo, lona, papel, etc. Podemos considerar gran formato a todo aquello que vaya más allá de un hoja A3.

2.3 El proceso de elección de un equipo de impresión

2.3.1 Lista de especificaciones técnicas

A continuación se listan las características técnicas de los equipos de impresión que pueden ser encontrados en los folletos de especificaciones provista por los distintos fabricantes. Dichas características son fundamentales para seleccionar qué modelo de equipo se necesita. Dependiendo del

¹²Extraído de <http://www.duiops.net/hardware/impresor/impresor.htm>

¹³http://www1.lexmark.com/es_MX/cartridges-ink-supplies/iso-page-yields/,
<http://isotc.iso.org/livelink/livelink/open/jtc1sc28>

uso que se le va a dar al equipo las distintas características van adquiriendo más o menos relevancia.

- Tipo (impresora / multifunción)
- Tecnología de impresión (laser / inyección de tinta / matriz de punto)
- Color
- Funciones (imprimir / escanear / copiar / envío de fax / envío digital)
- Volumen de impresión recomendado
- Conectividad (usb / puerto paralelo / red ethernet / wifi)
- Lenguajes de Impresión Soportados (GDI / PCL / PostScript / ESC/Page)
- Tipo de papel (normal, fotográfico)
- Tamaño de papel (A4, A5, etc)
- Velocidad de impresión
- Calidad de Impresión
- Cantidad de Memoria
- Duplex automático
- Pantalla táctil
- Impresión en A3
- Cantidad de hojas soportadas en la/las bandejas
- Adicionales (bandejas adicionales)

Otras características a tener en cuenta son:

- Período de garantía
- Precio de los insumos
- Servicio técnico local

2.3.2 Entrevista con el experto

En el *Apéndice A* se encuentra el contenido de los intercambios de mails con el experto (vendedor), que han sido dispuestos en forma de entrevista. Usando como guía lo manifestado por el experto hay cuatro características mínimas que considerar para un usuario común y son

uso (en qué ámbito será usada: oficina / hogar)
funciones (copiar, imprimir, etc)
color
volumen mensual
conectividad requerida

2.3.3 Heurísticas

Directamente de la “entrevista” con el experto extraemos las siguiente lista preliminares de reglas que luego deben ser validadas.

Si es para poco uso y una sola persona \Rightarrow necesita impresora chica

Si es para imprimir y escanear \Rightarrow necesita multifunción

Si es para uso corporativo \Rightarrow necesita impresora de Red

Si no quiere cables \Rightarrow necesita conexión wireless

Si es fanático de la tecnología \Rightarrow puede querer imprimir desde el iPod

Si es para altos volúmenes \Rightarrow el costo del toner es importante

Si el volumen mensual del equipo es menor a 2000 páginas \Rightarrow la impresora es chica

Si es para una oficina \Rightarrow necesita laser monocromo

Si es un estudio de arquitectura \Rightarrow necesita chorro de tinta para A4 o A3 o un plotter

Si es para el hogar \Rightarrow puede necesitar multifunción barata

Uso

Hay ciertos conceptos impresos como *impresora chica*, *alto volumen*, *poco uso*. En el contexto utilizado todos los términos se están refiriendo a la cantidad de páginas mensuales. Algo más con respecto a la palabra *uso* se la puede emplear para manifestar frecuencia de impresión, como para referirse a tipo de documentos a imprimir (fotos, textos, etc).

Otras posibles reglas extraídas de noticias y recomendaciones online podrían ser las siguientes.

Si se imprimen menos de 30 pag a color \Rightarrow cartucho único.

Si se imprimen más de 30 pag a color \Rightarrow cartuchos separados.

Si el uso es doméstico u ocasional \Rightarrow la velocidad no es relevante.

Si el uso es doméstico u ocasional \Rightarrow el tamaño de la bandeja no es relevante.

Si se deben imprimir muchas copias \Rightarrow la velocidad es relevante.

Si se quieren imprimir fotografías \Rightarrow no se recomienda una impresora domestica.

Si no se quieren imprimir fotografías \Rightarrow una resolución de 600 dpi alcanza.

Si el volumen de impresión es alto \Rightarrow se recomienda impresora láser y velocidad de impresión superior a 30 ppm

Table 2: Correspondencia entre volumen mensual y tamaño del grupo de trabajo

Tamaño del grupo	Ciclo mensual	Volumen
1 persona	10000 p./mes	hasta 1000 p.
1 a 3 usuarios	hasta 50000 p./mes	hasta 4000 p.
4 a 10 usuarios	hasta 60000 p./mes	hasta 5000 p.
más de 10 usuarios	hasta 85000 p./mes	hasta 8000 p.

Si es una empresa de publicidad \Rightarrow se recomienda laser color u otras tecnologías

Si es para formularios continuos \Rightarrow se recomienda matriz de punto

Si es para el hogar o pequeña oficina \Rightarrow inyección de tinta

Si es para el hogar o pequeña oficina y se necesita calidad fotográfica \Rightarrow inyección de tinta con 4 cartuchos.

Si el grupo de trabajo promedio es de 10 computadores \Rightarrow el tamaño de la bandeja de entrada es importante, y se prefiere de más de 200 hojas.

Si se requiere escanear en un estudio contable o de abogados \Rightarrow conviene chequear la superficie del escaner admite tamaño legal.

2.3.4 Perfil de usuario

Está claro que cuáles de todas las características que posee un equipo hay que tener en cuenta depende principalmente del uso que se le dará al equipo y en que ámbito se le dará ese uso (hogar, empresa, etc). Porque el uso determina varios del resto de las características.

2.3.5 Notas sobre volumen mensual

HP Argentina, en su página web¹⁴ clasifica sus productos en 5 categorías y guía al cliente que visita su página en busca de impresoras o multifunciones en esta clasificación de acuerdo al uso que tiene intenciones de darle el cliente.

En la página de Xerox¹⁵ dentro de los equipos ofrecidos para la pequeña y mediana empresa hay una clasificación de acuerdo al tamaño del grupo de trabajo. Dicha clasificación se compone de las categorías de *personal*, *grupo pequeño* (de 1 a 3 usuarios), *grupo mediano* (de 4 a 10), y *grupo grande de trabajo* (más de 10). Al revisar las especificaciones de los productos de las distintas categorías se observa una relación entre dichas categorías y el ciclo mensual de trabajo que está resumida en la tabla 2.

3. DISEÑO

La figura se presenta la una arquitectura de tres niveles¹⁶ del sistema.

¹⁴<http://www8.hp.com>

¹⁵<http://www.office.xerox.com/small-business/esuy.html>

¹⁶Una arquitectura de tres niveles es una especialización de la arquitectura cliente-servidor donde la carga se divide en tres partes (o capas) con un reparto claro de funciones: una capa para la presentación (interfaz de usuario), otra para el cálculo y otra para el almacenamiento (persistencia). Una capa solamente tiene relación con la siguiente.

Table 1: Volumen mensual según uso pretendido.

Volumen mensual	Uso pretendido
Hasta 500 páginas mensuales	Fotos y Documentos
Entre 500 y 1000 páginas mensuales	Impresión de documentos profesional
Entre 1000 y 5000 páginas mensuales	Empresa en crecimiento
Entre 5000 y 10000 páginas mensuales	Impresión de volumen para medianas empresas
Más de 10000 páginas mensuales	Soluciones para organizaciones Empresariales

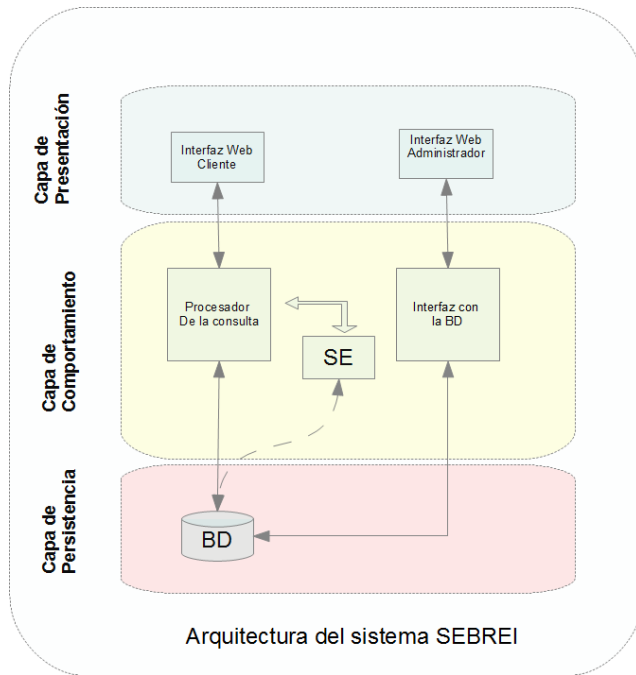


Figure 2:

3.1 Desiciones de Diseño

3.1.1 Acotación del dominio de tarea

El mercado de equipos de impresión es bastante grande. Por una cuestión de practicidad y tiempo el sistema desarrollado recomendará quiepos para los segmentos de mercados desde el hogar a empresas, pero siempre equipos que admitan imprimir en formato A4. Inicialmente el sistema no recomendará equipos para lo que se conocen como las artes gráficas, ni equipos que se encuentran dentro de la categoría de grandes formatos. De todas formas el diseño del sistema se plateará de forma tal que puedan ser incorporadas nuevas reglas que cubran estas categorías. También hay algunas características de los equipos e impresión que un determinado cliente pueda desear y que no sea necesario el asesoramiento de un profesional para darse cuenta si un equipo que le ofrecen sirve o no a sus necesidades. Un ejemplo de esta característica podría ser el tamaño físico del equipos, en este caso al cliente le basta mirar la lista de especificaiones técnicas y verificar si el tamaño del equipo que le ofrecen es el adecuado.

Si bien podría ser un adicional preguntar y recomendar teniendo en cuenta este tipo de cosas, no agrega inteligencia

al sistema y la mayoría de las veces las preguntas del sistema serán demasiado redundantes.

3.1.2 Representación del conocimiento

La principal tarea del sistema experto es determinar cuáles del conjunto de características que describen a un equipo son relevantes y que valores deben tener para satisfacer las necesidades del cliente. Una vez determinadas cuales son las características deseadas en un equipo se procede a ofrecer los modelos que cumplen los requerimientos.

Para determinar las características deseadas se usarán reglas de inferencia. Mientras que los distintos modelos serán hechos factuales en la base de conocimiento.

3.1.3 Determinación de las características requeridas

Para determinar algunas características será necesario preguntar al cliente, otras en cambio serán inferidas a partir de los datos provistos por el cliente.

Para el volumen mensual la clasificación de HP, define 5 conjuntos crips para el volumen mensual de impresión, se usarán dichos conjuntos para los valores del literal *uso*, los valores de la variable *uso* podrán ser *hogareño*, *profesional*, *pyme*, *pyme+* y *empresarial*.

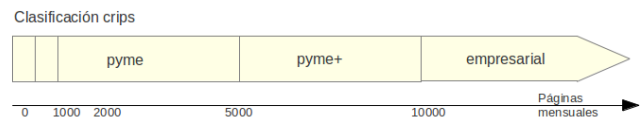


Figure 3: Clasificación según volumen de impresiones mensuales

Uso El uso vendrá determinado a través de consultar con el cliente cuantas páginas cree que en promedio imprimirá en un mes.

Aunque también se podría usar la aproximación que usa Xerox teniendo en cuenta el tamaño del grupo de trabajo estimar el volumen de páginas mensuales. De momento se usará la versión simplificada de preguntarle directamente al cliente.

3.2 Base de datos

Para permitir flexibilidad en la incorporación de nuevos equipos o modificación de los existentes, utilizamos un base de datos relacional para mantener las características de los equipos de impresión. La figura 4 muestra el diagrama de entidad relación de la base de datos.

4. IMPLEMENTACIÓN

Un primer prototipo se implemento enteramente en CLIPS. Antes de entrar en los detalles haré una breve introducción a clips.

Primeramente para arrancar CLIPS desde una shell de Linux se teclea clips y nos encontramos en un interprete de comandos. La notación es muy similar a la del lenguaje LISP.

(exit) Sale de CLIPS

(clear) Elimina datos y programas de la memoria.

(reset) Elimina la información dinámica de la memoria y resetea la agenda.

(run) Comienza la ejecución de un programa CLIPS.

4.1 Hechos y Reglas en Clips

En un su modo más sencillo, CLIPS opera manteniendo una lista de hechos (facts) y un conjunto de reglas (rules) con las que operar. Los hechos en CLIPS pueden ser de dos tipos: ordered facts y deftemplate facts. Un ejemplo de ordered facts es:

```
(color si)
(paginas 250)
```

Un ejemplo de deftemplate es:

```
(deftemplate questions
  (slot question (default ?NONE))
  (slot attribute (default ?NONE))
  (slot type (default ?NONE))
)
```

La creación de hechos se realiza mediante las sentencias assert (uno solo) o deffact (un conjunto), por ejemplo:

```
(assert(color si))
(assert (questions
  (question "¿Será necesario imprimir a color?")
  (attribute color)(type bool)
)
)

(deffacts mis-hechos
  (color no)
  (paginas 500)
  (questions
```

```
(question "¿Será necesario escanear?")
(attribute grupo) (type bool)
)
)
```

Las reglas en CLIPS están formadas por: Una parte izquierda (LHS) que define las condiciones a cumplir y una parte derecha (RHS) que define las acciones a realizar.

Sintaxis:

```
(defrule nombre-regla"comentario"
  (condicion-1)(condicion-2)...
=>
  (accion-1)(accion-2)...)
)
```

Para eliminar un hecho de la memoria de trabajo se usa el comando (retract f).

(printout) (read) (test) (operadores booleanes) (if)

4.2 Reglas

Para ayudar a visualizar el conjunto de reglas que componen el sistema ayuda a disponerlas en una estrura de árbol.

Una de las características principales a determinar es el volumen de páginas mensuales, con lo cual se seleccionará una de las categorías predefinidas para dicho característica. En la implementación se han usado las 5 categorías ya descritas para el volumen de páginas mensuales, pero ya se verá que el sistema admite cambiar dichos límites e incluso se pueden crear nuevas categorías desde la interfaz web administrativa.

El árbol correspondiente a la regla que determina la categoría uso que se necesita se encuentra en la figura 5.

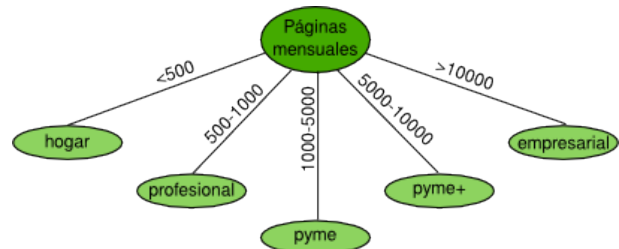


Figure 5:

Las reglas en clips quedan de la siguiente manera:

```
(defrule regla-uso1
  (paginas ?p)
  (test (< ?p 500) )
=>
  (assert (uso hogar)))
```

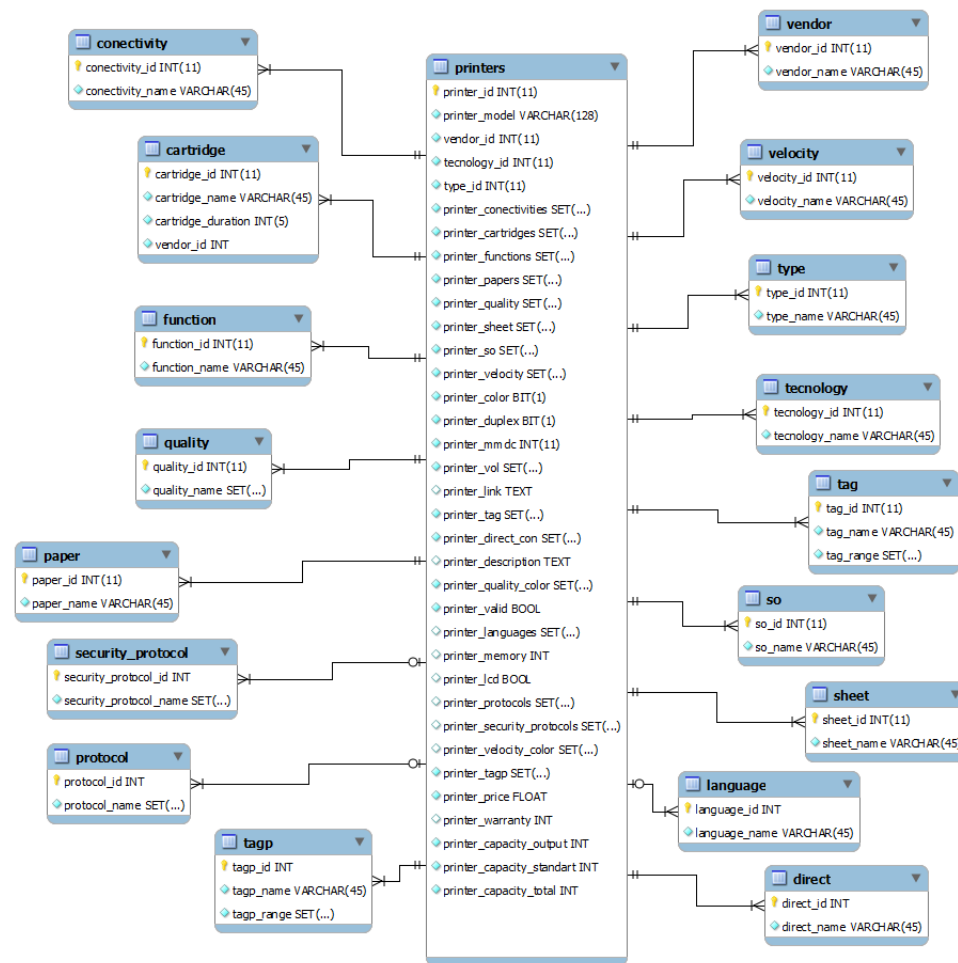


Figure 4: Modelo EER de la base de datos

```
(defrule regla-uso2
  (paginas ?p)
  (and (test (> ?p 500) )
        (test (<= ?p 1000)))
=>
  (assert (uso profesional)))

(test (> ?p 10000))
=>
  (assert (uso empresarial)))
```

```
(defrule regla-uso3
  (paginas ?p)
  (and (test (> ?p 1000))
        (test (<= ?p 5000)))
=>
  (assert (uso pyme)))
```

```
(defrule regla-uso4
  (paginas ?p)
  (and (test (> ?p 5000))
        (test (<= ?p 10000)))
=>
  (assert (uso pyme+)))
```

```
(defrule regla-uso5
  (paginas ?p)
```

5. DESARROLLO DE LAS INTERFACES WEB

5.1 Interfaz administrativa

La interfaz administrativa permite agregar nuevo equipamiento a la base de datos. Y cambiar parametros que afectan el comportamiento del sistema experto, como por ejemplo cambiar el límite de las categorías de volumen mensual de impresión o inclusive agregar nuevas.

5.1.1 Requerimientos

La interfaz web debe permitir al administrador del sistema agregar elementos a la base de datos y ajustar ciertos parámetros del SE. Permite:

- Logueo y autenticación
- Agregar/Modificar :
 - Impresoras

- conectividades
- marcas
- velocidades
- tecnologías
- tipo de equipos
- sistemas operativos
- tamaños de papel
- tipos de papel
- calidades de impresión
- funciones del equipo
- lenguajes de impresión
- consumibles
- Cambiar límites a los tags* (de cant de impresiones y de precios)
- usuarios
- Mantenimiento:
 - Actualizar tag a los equipos
 - Generar listas de impresoras
 - Actualizar reglas de uso

Los tags de precios se actualizarán mediante una función fuzzy cuyos parámetros pueden ser alterados por el administrador del sistema. Eso permitirá que los valores de la variable precio económica, moderado varien sus límites de acuerdo a la realidad económica del medio en donde se use el sistema.

5.2 Interfaz cliente

5.3 Testing

Agregar diseños de las interfaces web en WEBML

6. NOTAS SOBRE LA IMPLEMENTACIÓN

7. TESTING

Pruebas de las distintas componentes

8. HERRAMIENTAS UTILIZADAS

8.1 Selección de las herramientas

La herramientas necesarias para el desarrollo y documentación del sistema ha sido varias.

Clips

Como shell para sistemas expertos he seleccionado CLIPS por ser de dominio público, permite una rápida prototipación, cuenta con una muy buena documentación, tiene la capacidad de poder interoperar con otros lenguajes y la disponibilidad de varias extensiones y derivados.

Clipseg: Editor para clips

Base de datos

La base de datos está implementada en postgres. La idea inicial fue usar mysql, pero postgres cuenta con la posibilidad de definir tipos de datos propios y definir usar arreglos en la definición de las filas de una tabla lo cual resulta útil en este sistema. Workbench:

phppgadmin pgsql

Interfaces web

Para portar el sistema de manera online a través de un servidor web he usado la librería PyClips, la otra opción era PHLIPS ambas implementan un subconjunto de los comandos propios de CLIPS, pero PyClips implementa un número mayor de comandos y además cuenta con una muy buena documentación y un foro de discusión activo¹⁷.

PHP

La interfaz administrativa del sistema se desarrollo en PHP.

Webratio: WebML

Eclipse

Otras herramientas

clientes svn: tortoise (cliente svn para Windows), svn para linux

LibreOffice: Calc, Draw

Gimp: Editor de imágenes

Inkscape: editor de imágenes vectorial

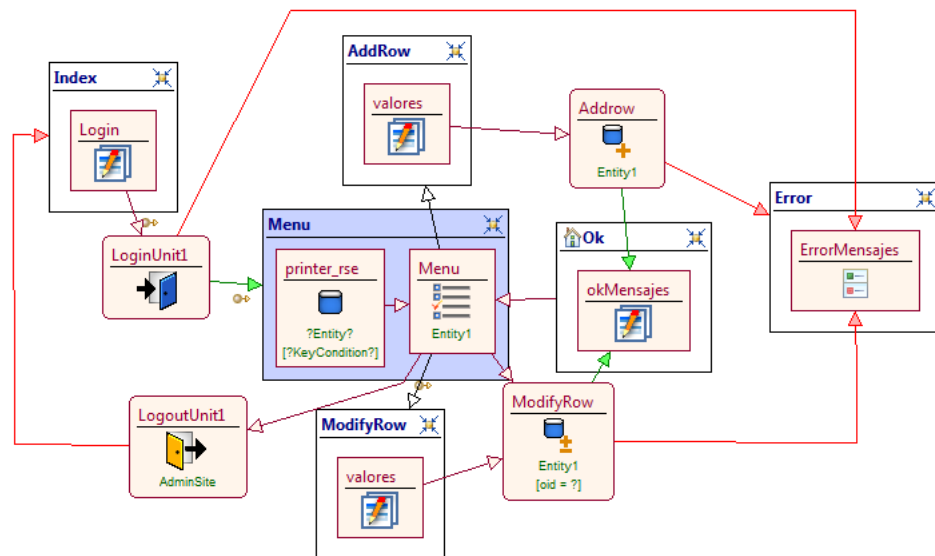
Editor de Latex

9. CONCLUSIÓN Y NOTAS FINALES

¹⁷<http://sourceforge.net/projects/pyclips/forums>

SEBREIweb_AdminSite

Erica Vidal



References

- [1] Otero Spagnuolo, M.
- [2] Fiorotto, A.
- [3] Montoya, M. A. *Impresoras, Curso online*.
- [4] Huergo, C.
- [5] 24734:2009, I. (2009).
- [6] Manual de Clips
- [7] Manual de pylclips

APPENDIX

A. ENTREVISTA CON EL EXPERTO

IC- ¿Qué tipo de información se le pide a un cliente cuando se acerca con intenciones de comprar un equipo de impresión?

Experto- Hay varias consideraciones cuando un interesado con consulta por impresoras, las relevantes son:

1. Qué tipo de uso le va a dar a la impresora, también el trabajo o profesión del cliente es un dato orientativo.
2. Producción de trabajo diario o mensual.
3. Si es para grupos de trabajo o para usar solo junto a una PC.
4. Si también escanean documentos, sacan fotocopias y/o usan fax

Con estos primeros datos ya tenemos para prepararle un abanico de soluciones de impresión. Es decir, si es para poco uso y una sola persona puede sugerirse una impresora chica, si también escanea se amplía a una multifunción. Si ya es corporativa conviene una impresora en RED, con print server interno. A veces el cliente para su hogar ya está pidiendo impresoras WIRELESS ya que no quiere cables dando vuelta por su casa, también hoy hay estos equipos. Además para los fanáticos de la tecnología tenemos las nuevas máquinas que imprimen desde un iPOD, etc. También hoy el asesoramiento va por el costo de los toner que al momento de la elección puede ser importante sobre todo para altos volúmenes.

IC- *¿Cuál debe ser aproximadamente el volumen de impresiones diario o mensual para que se considere que se necesita una impresora chica?, o sino ¿cuándo considero que una impresora es chica?*

Experto- *Normalmente el fabricante especifica en la folletería el volumen de trabajo, normalmente si el volumen mensual es de hasta 1.500 a 2.000 páginas por mes, en Argentina se estila comprar una máquina chica.*

IC- *Mencionaste que la profesión de un cliente es un dato orientativo para la recomendación de un equipo, en ese sentido necesitaríamos un poco más de detalle. ¿Qué tipos de profesiones indican qué tipos de equipos?*

Experto- *Por ejemplo, para una oficina lo más común es usar impresoras láser monocromo. Si es un estudio de arquitectura es probable que usen chorro de tinta ya sea formato pequeño o bien A3 (297 x 432 mm.) o bien plotter. Un diseñador gráfico o fotógrafo puede elegir una impresora HP Photosmart con 6 tintas de colores diferentes. Para el hogar una multifunción ink jet línea barata. Si entrás a la página de HP vas a encontrar que las impresoras las categorizan para:*

- *mercado hogareño*
- *mercado de diseño*
- *mercado corporativo*