

✓ **KELLY: UNA HERRAMIENTA PARA LA ADQUISICION DE CONOCIMIENTO HEURISTICO**

LUIS EDUARDO MUNERA

Matemático de la Universidad del Valle, Máster y doctor en informática de la Universidad politécnica de Madrid, Ex-profesor de la facultad de informática de la U.P.M. profesor del ICESI.

GRACE ALEXANDRA RITTER

Graduada en ingeniería de sistemas, ICESI.
Cargos ocupados: Ingeniero investigación y desarrollo, División Electrónica Carvajal S.A. asistente técnico centro de computo, ICESI.

JUAN M. MADRID

Estudiante de VII semestre de ingeniería de sistemas, ICESI.

1. INTRODUCCION

El cuello de botella en el desarrollo de sistemas expertos es la extracción o educción del conocimiento heurístico del experto por parte del ingeniero del conocimiento. Uno de los aspectos más difíciles de las tareas del ingeniero del conocimiento es ayudar al experto a estructurar el dominio del conocimiento, a identificar y formalizar los conceptos del dominio. Los expertos tienen tendencia a establecer sus conclusiones y el razonamiento que usan, en términos generales que son demasiado amplios para obtener un análisis

efectivo para la máquina. El que hace rápidamente juicios complejos sin reexaminar y reestablecer laboriosamente cada paso en su razonamiento.

Cuando un experto examina un problema, no puede articular fácilmente cada paso y puede incluso ignorar cada uno de los pasos individuales dados para alcanzar una solución. Puede achacarlo a la intuición o etiquetar como una repentinización lo que es el resultado de un proceso de razonamiento complejo basado en una gran cantidad de datos y experiencias recordadas. Este comportamiento del experto es lo

que Waterman (10) ha dado en llamar la "paradoja de la experiencia", esto es, que los expertos más competentes, son los menos hábiles para describir el conocimiento que usan para resolver los problemas. Incluso peor, pues estudios recientes han demostrado que cuando los expertos intentan explicar cómo alcanzan una conclusión, con frecuencia constituyen líneas plausibles de razonamiento que se parecen poco a su actividad real para solucionar problemas. Este efecto tiene, al menos, dos implicaciones importantes para la construcción de sistemas expertos.

1. Los expertos necesitan ayuda externa para clasificar y explicar su pensamiento y la forma en que resuelven sus problemas. Esto de paso recomienda que el experto y el ingeniero del conocimiento no sean la misma persona.
2. Los ingenieros del conocimiento no deben creer todo lo que dicen los expertos. Los ingenieros del conocimiento deberán desarrollar hipótesis de trabajo basadas en lo que dicen los expertos y verificarán las hipótesis y su consistencia, proponiéndoles los expertos que resuelvan nuevos problemas que necesitan el conocimiento hipotetizado.

2. TEORIA DE LA CONSTRUCCION PERSONAL

Las dificultades señaladas anteriormente son bien conocidas en psicología, por lo que algunas escuelas han desarrollado técnicas haciendo uso de una interacción verbal, a través de técnicas de entrevistas que vencen las defensas cognitivas. Una de dichas técnicas es la "teoría de la construcción personal" del psicólogo George Kelly, que está siendo aplicada con éxito en la ingeniería del conocimiento.

La teoría de la construcción personal, formulada por Kelly (8), asume que la gente típicamente usa dimensiones cognitivas llamadas "construcciones" para evaluar su experiencia.

Las construcciones son "patrones" que una persona crea en su intento de aproximarse a las realidades que componen el mundo.

Kelly propone que toda actividad humana debe ser vista como un proceso de anticipación del futuro construyendo la replicación de eventos, y revisando las construcciones de acuerdo con el éxito de su participación.

Kelly desarrolló su teoría en el contexto la psicología clínica y por lo tanto concernía a tener técnicas para vencer las defensas cognitivas y extraer las construcciones subyacentes en el comportamiento. En un contexto no clínico este es el problema de la ingeniería del conocimiento.

Kelly teoriza que cada individuo se comporta como un "científico personal", que predice y controla el entorno, formando teorías e hipótesis, y validándolas por medio de una categorización y clasificación de las experiencias.

Su modelo implica que las personas predicen y controlan eventos, formando hipótesis relevantes y luego contrastándolas con la evidencia disponible. Esas hipótesis son derivadas de las relaciones entre las construcciones que articulan la estructura "lógica" de un sistema de construcciones personales individuales.

3. EL EMPARRILLADO

Para ayudar a la gente a explorar sus propios sistemas de clasificación, Kelly desarrolló una herramienta psicoanalítica llamada "The repertory grid test".

El emparrillado (repertory grid) de Kelly es una vía de representar las construcciones personales como un conjunto de distinciones hechas acerca de elementos relevantes al dominio del problema. En psicología clínica este dominio a menudo serán las relaciones personales y los elementos pueden ser

miembros de la familia y amigos. En el desarrollo de sistemas expertos los elementos serán entidades claves.

Un emparrillado es esencialmente un complejo test de clasificación u ordenamiento en el cual una lista de elementos son juzgados sucesivamente sobre la base de un conjunto de construcciones bipolares. Los elementos pueden ser entidades abstractas o concretas, y se puede decir que definen operacionalmente el universo del discurso del emparrillado. Las construcciones, por definición, representan distinciones bipolares. Por ejemplo, una madre puede usar la construcción "excitado/calmado", entre otros, para interpretar el comportamiento de su hijo. Las construcciones representan las formas bajo las cuales son juzgados los elementos en cuanto a diferencias y semejanzas con respecto a otros. Así que un emparrillado puede ser considerado como un mapa de los elementos sobre las construcciones. Los datos generados por cada sujeto son introducidos en una tabla o parrilla, en la cual

hay una columna para cada elemento y una fila para cada construcción. Cada intersección de la parrilla contiene un número indicando cómo una construcción dada fue aplicada a un elemento particular.

En dominios de aplicación tales como la radiología, los elementos frecuentemente representan enfermedades u otros estados patológicos, mientras que las construcciones delinean síntomas (u otra información clínica), datos físicos o de laboratorio o evaluaciones de un test.

Un ejemplo de emparrillado es el siguiente de GAINES (7), cuyo propósito es la valoración de personal y consta de 10 elementos (en este caso miembros del personal) y de 14 construcciones bipolares. La escala de valoración empleada es de 1 a 5. Se usa para hacer más precisa la distinción en situaciones donde no se desea localizar definitivamente un elemento en un polo de construcción. Un valor de 1 significa el polo de la izquierda y un valor de 5 significa el polo a la derecha.

	1	2	3	4	5	6	7	8	9	10	
INTELIGENTE	1	1	4	5	3	3	5	2	3	5	Torpe
VOLUNTARIOSO	1	2	4	5	1	1	4	3	1	2	No voluntarioso
NOVATO	1	2	3	5	4	4	4	1	4	3	Experimentado
REQUIERE POCA SUPERVISION	3	1	4	5	2	1	5	2	2	3	Requiere mucha sup.
MOTIVADO	1	1	4	5	2	2	5	3	3	2	Desmotivado
ASEQUIBLE	3	2	2	5	1	1	5	1	2	3	No asequible
PACIFICO	3	4	5	2	2	3	1	5	4	5	Agresivo
SOÑADOR	1	1	5	4	2	3	1	3	4	4	No soñador
EMPRENDEDOR	2	1	5	5	1	3	5	3	4	5	No emprendedor
CREATIVO	1	1	5	5	2	3	4	3	4	5	No creativo
SERVICIAL	4	3	4	2	3	5	1	4	5	5	No servicial
PROFESIONAL	1	2	3	3	2	1	5	2	4	4	Menos profesional
MUCHA AUTOESTIMA	2	1	3	4	1	2	5	2	3	4	Poca autoestima
CONFUSO	2	2	5	4	3	5	1	5	3	1	Ordenado

4. ANALISIS DEL EMPARRILLADO

Originalmente Kelly usó un método de análisis factorial no-paramétrico cuando analizó el emparillado. Los resultados ayudaron a Kelly y a sus pacientes a entender el grado de similitud entre los rasgos. El llamó un rasgo al opuesto de una construcción, e hipotetizó que cada construcción representa algún concepto interno para el paciente. En los emparillados de Kelly se utilizaron únicamente dos valores: "1" "0". El 1 significa que el predicado que representaba el polo izquierdo de una construcción es "verdadero" para ese elemento. Recíprocamente, la asignación de un 0 indica que el predicado que representa el polo izquierdo de la construcción es "falso", mientras que el predicado que representa el correspondiente polo derecho es "verdadero". Consecuentemente tales emparillados binarios pueden ser vistos como matrices booleanas de valores de verdad.

En la actualidad el emparillado es una representación de la visión del experto de un problema particular. Es una herramienta útil en la educación de conocimiento.

Actualmente la valorización se hace sobre una escala lineal de 1 a n, en donde 1 es una asignación al polo izquierdo de la construcción, y n es una asignación al polo derecho. Los números en el rango (2 ----> n-1) son empleados para representar categorías intermedias. Para permitir un punto medio natural, N es usualmente un pequeño número impar como 3,5, ó 7.

Existen varias aproximaciones al análisis del emparillado, como por ejemplo el análisis factorial no paramétrico empleado por Kelly, o el multidimensional scaling, o el cluster análisis, etc.

La aproximación que nos interesa se debe a Gaines y Shaw (7). Ellos proponen una semántica difu-

sa basada en conjuntos difusos en la que si un elemento E tiene un valor 0 sobre la construcción C donde el rango de valores es de 1 a N, entonces decimos que E tiene un grado de pertenencia de $(v-1)/(n-1)$ al polo derecho de la construcción C y un grado de pertenencia $(n-v)/(n-1)$ al polo izquierdo de la construcción.

Podemos considerar cada polo como representando un predicado lógico que define un conjunto difuso y el emparillado como un conjunto de conjuntos difusos, uno para cada predicado. Estos conjuntos difusos son entonces usados para la derivación de "supuestos" entre construcciones basados en la lógica difusa de Lukasiewicz L1. Así que si, X es el conjunto difuso definido por el predicado X y el grado de pertenencia del elemento E al conjunto difuso X es XE, entonces la media de similitud adoptada es:

$X \longleftrightarrow Y \approx \square (X \approx Y)$ que en la forma de Lukasiewicz sería:

$$\square (X \approx Y) = \min_e (1 - \tilde{X}_e + \tilde{Y}_e, 1 - \tilde{Y}_e + \tilde{X}_e) (*)$$

Esta formulación es interesante porque extiende la relación de equivalencia simétrica usada en clustering a la relación de suposición (implicación) asimétrica usada para inferencia en sistemas expertos.

$$X \rightarrow Y \approx \square (X > Y) = \min_e (1, 1 - \tilde{X}_e + \tilde{Y}_e) (**)$$

\square es el operador de necesidad de la lógica modal.

Como ejemplo consideremos un sistema experto que ayuda a un ingeniero del software acerca de qué manejador de bases de datos usar para una aplicación específica. Supongamos que se ha educido el siguiente emparillado:

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
(1)	5	5	3	5	5	5	5	3	5	5	5	5	5	5	1	1
(2)	5	2	1	5	5	5	1	5	5	5	5	5	1	5	3	3
(3)	5	5	5	5	5	5	5	5	5	1	1	1	5	1	5	5
(4)	5	5	5	5	5	1	5	5	1	1	1	1	5	5	5	5
(5)	1	1	1	1	1	1	1	1	5	1	1	5	1	1	1	1
(6)	5	1	5	5	3	5	5	5	1	1	5	5	5	1	5	5
(7)	3	2	2	4	3	3	3	1	3	5	2	3	3	4	1	5
(8)	5	3	3	4	4	2	3	4	3	3	1	1	3	4	1	5
(9)	5	5	5	1	1	5	5	2	5	5	5	5	3	2	5	5
(10)	1	2	5	3	3	5	3	3	5	5	5	5	5	1	5	5

1. ims 2. system-2000 3.inquiere 4.idms 5.total 6.sql/ds 7.model-284 8. ramis 9. dms-170 10. rim 11. oracle 12. ingres 13. adabas 14. sir 15. easytrieve 16. creatabase

- (1) Recuperación de texto/No recuperación de texto.
- (2) Archivos invertidos/No archivos invertidos
- (3) Corre sobre Vax/No corre sobre Vax
- (4) Relacional/No relacional
- (5) Corre sobre Ibm Mainframe/No corre sobre Ibm Mainframe
- (6) Corre sobre Cdm Mainframe/No corre sobre Cdm Mainframe
- (7) Generador de informes/No generador de informes.
- (8) Requiere bajo nivel de experiencia/Requiere alto nivel de experiencia.
- (9) Red/No red
- (10) Jerárquico/No jerárquico.

Los conjuntos difusos definidos por el emparillado son:

$$1 = \langle \emptyset, \emptyset, 5, \emptyset, \emptyset, \emptyset, \emptyset, 5, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, 1, 1, \rangle$$

$$2 = \langle \emptyset, 75, 1, \emptyset, \emptyset, 1, \emptyset, \emptyset, \emptyset, \emptyset, 1, \emptyset, \emptyset, 5, \emptyset, 5 \rangle$$

$$3 = \langle \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, \emptyset, 1, 1, 1, \emptyset, 1, \emptyset, \emptyset \rangle$$

$$4 = \langle \emptyset, \emptyset, \emptyset, 1, \emptyset, \emptyset, 1, 1, 1, \emptyset, \emptyset, \emptyset, \emptyset \rangle$$

$$5 = \langle 1, 1, 1, 1, 1, 1, 1, \emptyset, 1, 1, \emptyset, 1, 1, 1, 1 \rangle$$

$$6 = \langle \emptyset, 1, \emptyset, \emptyset, 5, \emptyset, \emptyset, \emptyset, 1, 1, \emptyset, \emptyset, 1, \emptyset, \emptyset, \rangle$$

$$7 = \langle \emptyset, 5, \emptyset, 75, \emptyset, 75, \emptyset, 25, \emptyset, 5, \emptyset, 5, \emptyset, 5, 1, \emptyset, 5, \emptyset, 75, \emptyset, 5, \emptyset, 5, \emptyset, 25, 1, \emptyset \rangle$$

8 = <0,0.5,0.5,0.25,0.25,0.75,0.5,0.25,0.5,0.5,1,1,0.5,0.25,1,0>

9 = <0,0,0,1,1,0,0,0.75,0,0,0,0.5,0.75,0,0>

10 = <1,0.75,0,0.5,0.5,0,0.5,0.5,0,0,0,0,1,0,0>

y sus opuestos 1,2,3,4,5,6,7,8,9,10, en donde por ejemplo:

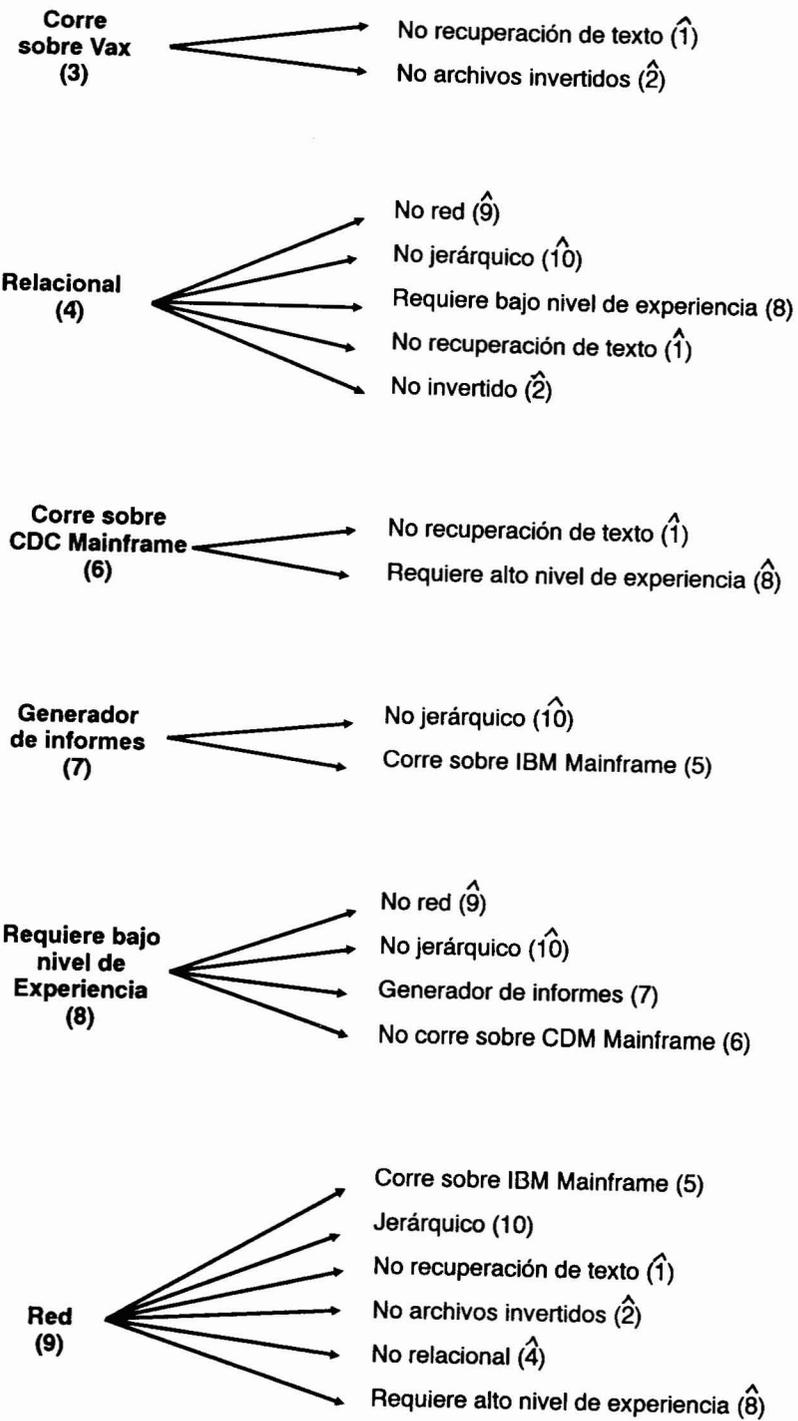
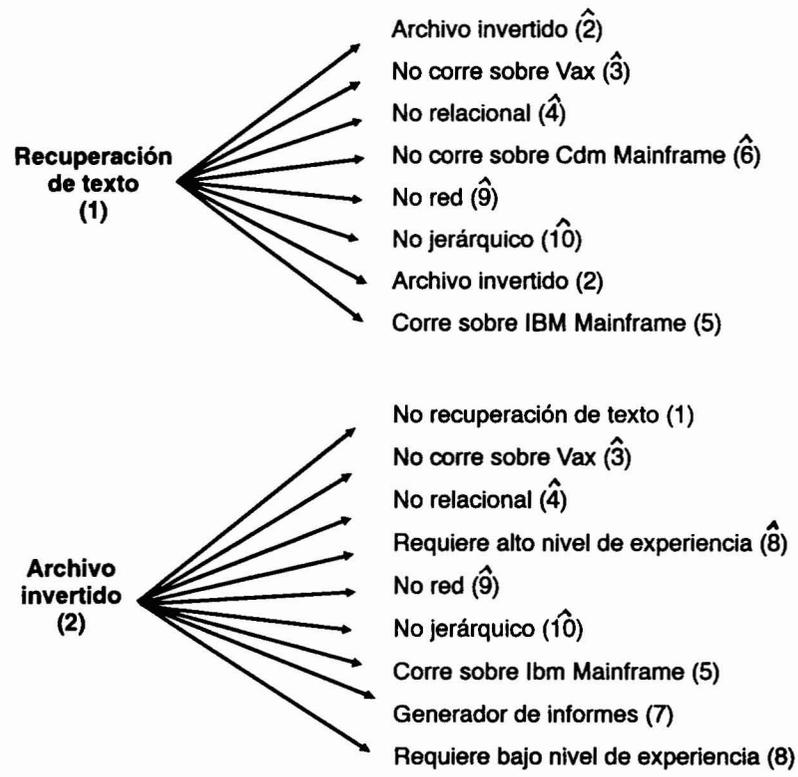
$\hat{1}$ = <1,1,0.5,1,1,1,1,0.5,1,1,1,1,1,1,0,0>

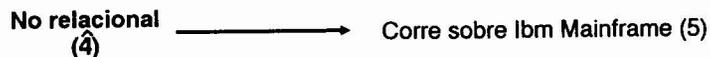
Utilizando la fórmula (***) obtenemos las implicaciones hipotéticas, por ejemplo:

1 \rightarrow 2: Min 1,1,1,1,1,1,1,0.5,1,1,1,1,1,0.5,0.5=0.5

Si suprimimos las suposiciones con factor de certeza menor de 0.5, obtenemos:

1 \rightarrow 2, 5, $\hat{2}$, $\hat{3}$, $\hat{4}$, $\hat{6}$, $\hat{9}$, $\hat{10}$ 2 \rightarrow 5, 7, 8, $\hat{3}$, $\hat{4}$, $\hat{8}$, $\hat{9}$, $\hat{10}$
 4 \rightarrow 8, $\hat{9}$, $\hat{10}$ 6 \rightarrow $\hat{8}$ 7 \rightarrow 5, $\hat{10}$, 8-7, $\hat{9}$, $\hat{10}$
 9 \rightarrow 5, 10, 10 \rightarrow 5 $\hat{4}$ \rightarrow 5 $\hat{5}$ \rightarrow 7, 8





5. GENERACION DE REGLAS

Varias herramientas de adquisición del conocimiento han sido desarrolladas, algunas de las cuales emplean lógica difusa en el análisis del emparrillado. por ejemplo Entail (Gaines y Shaw [7]) o ETS (Boose [2]), etc.

Estas herramientas actúan como sistemas "front end" a herramientas "shell" para el desarrollo de sistemas expertos.

La idea es sencillamente convertir las "implicaciones hipotéticas en reglas y trasladar estas al shell. Veamos por ejemplo el caso de Ets (Expert Transfer

System) desarrollada por Boose (2). La herramienta Ets genera un grafo de supuestos como el que generamos para el sistema de ayuda en la escogencia de un manejador de bases de datos. Luego Ets genera dos tipos de reglas de producción heurísticas: reglas conclusión (reglas Backward) y reglas intermedias (reglas forward).

Cada regla es generada con un factor de certeza y una vez generadas, todas las reglas pueden ser revisadas y modificadas por el experto. Las reglas conclusión son creadas a partir de los valores del emparrillado. Cada valor tiene el potencial para generar una regla. El experto es interrogado primero

para que cuantifique la importancia de cada construcción en términos de su potencial importancia en solucionar un problema. Entonces mediante un algoritmo que toma en cuenta los valores del emparrillado, la importancia relativa de la construcción y mediante una combinación de factores de certeza se puede generar una regla conclusión.

En el caso de Kelly V.1.0, generamos directamente a partir de la parrilla y de las ponderaciones dadas por el experto, los dos tipos de reglas, las reglas conclusión (utilizadas en un razonamiento hacia atrás) y las reglas antecedente (utilizadas en un razonamiento hacia adelante).

La herramienta no genera un grafo de supuestos como lo hace ETS, sino que genera directamente las reglas para que sean validadas inmediatamente por el experto con la ayuda de un motor de inferencias que posee la herramienta y que permite la creación de prototipos rápidos.

6. OBTENCION DEL EMPARRILLADO

Antes de diseñar una parrilla es importante definir el problema sometido a investigación. El experto, o el ingeniero del conocimiento, debe establecer un objetivo claro para su análisis. Problemas adecuados pueden ser: la clasificación de enfermedades de la piel, evaluación de candidatos a un trabajo, etc. Estos deberían ser específicos antes que generales. Un problema más general y más difícil de definir debería ser algo así como las características de

un buen diseño, o la clasificación de enfermedades.

Después de haber definido su objetivo el experto establece la parrilla acerca del diseño de la tarea. Si él puede escribir los elementos y construcciones directamente, entonces la parrilla puede ser analizada sin embargo, en la mayoría de los casos la educación no es trivial e implica repetidas comparaciones de elementos. Es habitual para el experto producir primero sus elementos. Luego se comparan y contrastan con el fin de obtener más detalles de posteriores elementos y construcciones. Un grupo de tres es el menor a partir del cual se pueden describir una similitud y una diferencia. Esta comparación puede ser la base para una nueva construcción los grupos de tres pueden elegirse aleatoria o sistemáticamente. Así por ejemplo:

Ic: Considere los elementos Pepe, Juan y Paco; elegir los dos más similares.

E: Pepe y Paco son similares y diferentes de Juan.

Ic: ¿Qué hace que Pepe y Paco sean similares?

E: Son más maduros.

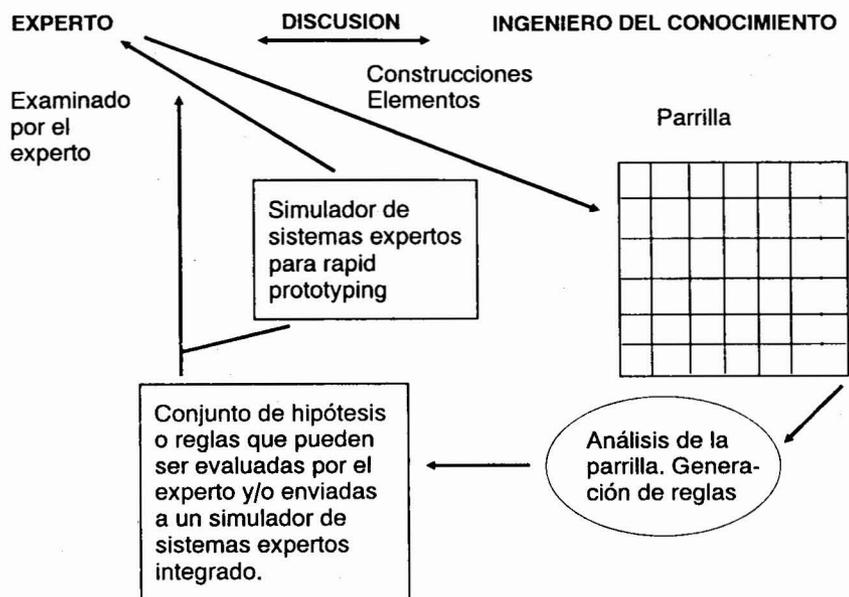
Ic: ¿Qué es lo opuesto a maduros que describe Juan?

E: Es inmaduro.

Ic: Los polos de la construcción son maduro e inmaduro. con 1 para inmaduro y 5 para maduro.

7. CICLO DE LA ADQUISICION DEL CONOCIMIENTO

La figura resume el ciclo de adquisición del conocimiento basado en el emparillado.



Kelly V.1.0 permite automatizar el ciclo de adquisición del conocimiento, mediante interfases amigables que utilizan iconos, ventanas, scroles, etc.

La herramienta de facilidades de edición, almacenamiento y recuperación de información.

También permite la creación y manipulación de parrillas no sólo para un experto, sino para varios expertos.

La herramienta guía el proceso de construcción en la parrilla, señalando la definición de elementos, atributos, opuestos, valorizaciones y ponderaciones que deben realizar los expertos. Luego de manera automática genera las reglas conclusión y antecedente con sus respectivos factores de certeza.

La herramienta posee un modo de consulta y simulación, distinto al modo

de desarrollo y que permite la creación muy rápida de prototipos que muestran la dinámica inferencial generada por las reglas.

Este modo de simulación consta de un motor de inferencias que realiza un encadenamiento mixto de las reglas conclusión y antecedente.

La herramienta permite la intervención de varios expertos en la solución de un mismo problema, con lo cual permite la creación de varias parrillas, una para cada experto y además es capaz de generar una parrilla de consenso cuando existen discrepancias entre los expertos.

REFERENCIAS

1. BENNETT, J.J. "Roget: Acquiring the conceptual structure of Diagnostic Expert System", *IEEE proceedings workshop on principles of knowledge based systems*, Denver Colora-

do, diciembre 1984, pp. 83-88. (1984).

2. BOOSE, J.H. "A knowledge acquisition program for expert systems based on personal construct psychology". *International journal of Man-machine studies*, Vol. 23, pp. 495-525. (1985).

3. BRUXELLES J. RACCAH P.Y. "Information and argumentation: How to express consequences". *Cognitiva 87*, May 1987, París, La Villete. (1987).

4. DIENG R., TROUSSE B. "3DKAT, a dependency-driven dynamic knowledge acquisition tool", *Proceedings of third international symposium of knowledge engineering*, Madrid, octubre 1988, pp. 85-93. (1988).

5. DUCROT O. *Les échelles argumentatives* París, 1980. Eds. de minuit. (1980).

6. FORD K.M., y otros "An approach to knowledge acquisition based on the structure of personal construct systems". *IEEE Transactions on knowledge and data engineering*, Vol. 3, Nº 1, Marzo 1991, pp. 78-87. (1991).

7. GAINES B.R., Show J.L.G. "Induction of Inference rules for expert systems". *Fuzzy sets and systems*, Vol. 18, pp. 315-328. (1986).

8. KELLY G.A. *The psychology of personal constructs*, New York, Norton. (1955)

9. KOSKO B. Fuzzy cognitive maps; *International journal of Man-Machine studies*, Vol. 24, Jan. 1986, pp. 65-75. (1986).

10. WATERMAN A *guide to expert systems*. Addison-Wesley, New York, (1986).