

El SETDC: un sistema basado en el modelo TDC.

Lic. José E. Medina Pagola  
Profesor Auxiliar

Ing. Giselle Castro Dorticós  
Profesor Instructor

Ing. Lisset Alfaro Remesal  
Profesor Instructor

Departamento de Informática  
Facultad de Ingeniería Industrial  
Instituto Superior Politécnico José Antonio Echevarría  
Carretera Central Manuel Martínez Prieto  
Marianao. Ciudad de la Habana, Cuba.

## El SETDC: un sistema basado en el modelo TDC.

Los elementos esenciales de los sistemas expertos (SE) son la base de conocimientos y la máquina de inferencias. A través de ellos los ingenieros en sistemas pueden desarrollar aplicaciones en las que las informaciones se representen de una forma más declarativa y en las que sean posibles inferir nuevas informaciones a partir del conocimiento representado.

Los SE están compuestos no solo de los elementos anteriores, sino que también pueden incluir, a partir de las secuencias deductivas desarrolladas, la habilidad de explicar el razonamiento seguido que justifica las soluciones tenidas, de una forma cercana al razonamiento que seguirían los expertos de esa actividad.

Las características antes mencionadas, unido a las posibilidades de representar el conocimiento mediante una lógica polivalente, hacen posible el desarrollo de aplicaciones en variados campos, tales como:

- el diagnóstico de fallas
- el pronóstico de diversos procesos
- la organización y planificación inteligentes de múltiples actividades
- la enseñanza o tutoría inteligente
- etc

El SETDC es un sistema vacío que permite desarrollar diversas aplicaciones afines con los SE utilizando como esquema de diseño al modelo TDC.

En este trabajo se exponen las características fundamentales del modelo TDC y su concreción en este sistema, así como algunas consideraciones en su aplicación y su desarrollo ulterior.

### El modelo TDC

El modelo TDC (de Tablas de Decisión Categorical) es un esquema conceptual para la representación y manipulación del conocimiento. (1 y 2)

En su concepción este modelo incluye elementos de otros, tales como el cálculo de predicados, los basados en reglas de producción y las redes semánticas. Sin embargo, el modelo TDC no se reduce a ellos ya que este integra a aquellos de una forma propia, enriqueciéndolos con nuevas estructuras y mecanismos.

La base de conocimientos en este modelo es representada mediante tres elementos básicos:

- la red taxonómica o jerárquica de categorías
- el conjunto de hechos
- las tablas de decisión

Todos estos elementos se integran en un mecanismo inferencial de backward chaining perfeccionado con procesos específicos de case y procesamiento conjunto de antecedentes de tablas de decisión, denominado también como proceso alfa.

Estos elementos y otros específicos serán detallados al describir el sistema SETDC.

### El SETDC 1.0

El SETDC 1.0 es la primera versión realizada sobre el modelo TDC (3).

Este sistema está conformado por los siguientes subsistemas:

- el generador de la base
- el monitor de la base
- los utilitarios del sistema

El generador (GENTDC) es un programa capaz de procesar la descripción de una base de conocimientos editada como un fichero ASCII; analizar sintácticamente dicha descripción y generar la base de conocimientos que será efectivamente procesada por la aplicación.

El monitor de la base (MONTDC) permite verificar el correcto procesamiento de la base generada por GENTDC. Mediante este programa no solo se pueden analizar las inferencias a obtener y el funcionamiento de cada utilitario del sistema, sino que también puede comprobarse paso por paso cómo se efectúa el proceso inferencial con la base dada.

Los utilitarios del SETDC son un conjunto de procedimientos que desarrollan todas las facilidades del sistema. Estos están organizados en units del Turbo Pascal 5.5.

De lo anterior se infiere que las aplicaciones finales pueden programarse en dicho lenguaje, asumiendo las interfaces deseadas y garantizando diseños con toda la profesionalidad posible.

Para el empleo de estos subsistemas debe seguirse los pasos que se indican en el siguiente esquema:

En este esquema no solo se indicará la interrelación entre estos tres subsistemas, sino que además se expresa el carácter de prototipo y de perfeccionamientos sucesivos (a) que está presente en la creación de todo sistema informativo (4) y de todo SE (5).

#### La descripción de la base de conocimientos

Como fue mencionado anteriormente, la base de conocimientos es descrita mediante cualquier editor de texto utilizando un lenguaje específico.

En dicho lenguaje se reconocen 4 partes denominadas también como bases, las cuales son:

- la base de átomos.
- la base de predicados.
- la base de hechos.
- la base de reglas.

En la base de átomos se definen todas las categorías atómicas a ser utilizadas en la descripción del conocimiento, así como las relaciones que se establecen entre ellas.

Un átomo es una categoría simbólica que denota algún concepto y que puede relacionarse con otras categorías mediante relaciones de ordenamiento, denominada en el modelo TDC como relaciones de afinidad. Por ejemplo, si se definiera la siguiente base de átomos:

ATOMOS

raúl:talentoso;

raúl:alumno\_instructor;

talentoso, alumno\_instructor:alumno;

centro\_de\_investigación:centro;

se estarían definiendo seis categorías atómicas con cinco relaciones de afinidad explícitas entre ellas.

Tales relaciones de afinidad pueden definirse no solo en forma arbórea, como puede apreciarse en el ejemplo, sino también en forma de grafo, estableciéndose una red taxonómica o jerárquica entre las categorías atómicas.

Esta característica posibilita su manipulación a través de un proceso de herencia múltiple, permitiendo definir intersecciones o conjunciones semánticas de los conocimientos heredados.

En el modelo TDC, además de las categorías atómicas, se tienen las categorías de objetales, aunque en esta primera versión estas últimas no se han incluido.

En la base de predicados se definen todos los predicados que utiliza la base de conocimientos. En la definición de cada predicado se indica no solo los argumentos que estos permiten, sino también las categorías atómicas superiores en ordenamiento admitidas por esos argumentos.

A modo de ejemplo supóngase la siguiente base de predicados:

#### PREDICADOS

responsable(alumno);

puede\_trabajar(alumno,centro);

.

En esta base se definen dos predicados, cada uno con uno y dos argumentos y tales átomos como categorías superiores permitidas para esos argumentos.

Además de las informaciones antes descritas, cada predicado admite la definición de diversos tipos, los cuales serán analizados junto a los procesos inferenciales incluidos en el sistema.

La base de hechos define todas las informaciones incondicionalmente ciertas o falsas que servirán como hechos de partida en el SE que se diseñe.

Los argumentos permitidos en los hechos pueden ser constantes atómicas o los tipos Pascal longint, real, boolean y string, garantizando con ello una mejor compatibilidad con los procesos a diseñar en Turbo Pascal.

Las constantes atómicas pueden cuantificarse universal (V) o existencialmente (E) de la siguiente forma:

$p(a)$  se corresponde con: V a  $p(a)$

$p(a^{\wedge})$  se corresponde con: E a  $p(a)$

no  $p(a)$  se corresponde con: V a  $\sim p(a)$

no  $p(a^{\wedge})$  se corresponde con: E a  $\sim p(a)$

siendo "a" un átomo cualquiera.

Nótese que la cuantificación de un término atómico se realiza en el propio argumento, lo cual, aunque simplifica su notación, dificulta la modificación del orden de cuantificación entre los términos de un mismo predicado. Esta limitación, además de ser soluble, no repercute significativamente en los SE que se realicen, pues es comprensible que el uso de la cuantificación existencial resulta poco frecuente.

Tanto la base de hechos como la de reglas permiten definir valores de certezas asociados con las informaciones contenidas en dichas bases.

Estos valores de certezas se toman en el intervalo  $[-1.0, 1.0]$ , considerando sus extremos lo absolutamente falso y cierto y su centro (0.0) lo absolutamente desconocido.

A modo de ejemplo supóngase la siguiente base de hechos:

HECHOS

responsable(alumno\_instructor)/0.8;

En esta base se ha definido un único hecho, el cual expresa con una certeza de 0.8 que todos los alumnos instructores son alumnos responsables.

En la base de reglas se definen las reglas de producción agrupadas en tablas de decisión (TD).

Según el modelo TDC, una TD contiene los siguientes elementos:

	Reglas
Condiciones	Estados condicionales
	Certeza
Acciones	Estados de las acciones

Siendo las condiciones y acciones de la tabla los antecedentes y consecuentes de las reglas y los estados de cada una de ellas la forma en que se incorporan tales elementos para cada regla.

En el modelo TDC se permiten dos formatos de estados condicionales, correspondiéndose con las tablas de entradas limitadas y ampliadas. En esta versión del SETDC se permiten diseñar solamente TD de entrada limitada.

Una TD de entrada limitada es aquella que admite como estados condicionales tres categorías básicas: lo cierto (si), lo falso (no) y lo desconocido (\*), además de permitir cualquier combinación de tales estados.

Las categorías anteriores establecen los valores de certezas aceptados por cada regla para las inferencias asociadas con cada antecedente. Para ello, tales valores de certezas son clasificados en tres subintervalos de la siguiente forma:

$\text{umbral} < \text{certeza} \leq 1.0$  para lo cierto

$-\text{umbral} \leq \text{certeza} \leq \text{umbral}$  para lo desconocido

$-1.0 \leq \text{certeza} \leq -\text{umbral}$  para lo falso

En donde el valor de umbral es un valor modificable por el usuario y asumido de forma estándar como 0.2.

Al igual que los hechos, los antecedentes y consecuentes de tablas admiten cualquier constante de las asumidas en el sistema como válidas. Sin embargo, en los predicados definidos como antecedentes o

consecuentes también se permiten argumentos variables.

Un argumento variable sigue el formato:

a # N

donde:

a: es un átomo cualquiera que define la categoría superior admitida por ese argumento.

N: es un entero positivo o cero.

Por ejemplo, si se definiera la siguiente base de reglas:

REGLAS

TABLA ejemplo

responsable(talentoso#) |si;

CERTEZA |0.9;

puede trabajar(talentoso#,centro\_de\_investigación)|\*;

entonces se habría definido una TD de una sola regla, la cual indica, con una certeza de 0.9, que aquel alumno talentoso que sea responsable puede trabajar en cualquier centro de investigación, por lo que se pudiera inferir, con la base de átomos y de hechos ejemplificada antes, que Raúl puede trabajar en tales centros con cierta certeza.

### El proceso inferencial

En el SETDC 1.0 los predicados a ser verificados inicialmente son los antecedentes de alguna tabla que se lanza como TD raíz del proceso. Las proposiciones inferidas a través de los consecuentes de la tabla raíz son almacenadas en bases de hechos definidas como de inferencias. Los antecedentes de la tabla raíz son verificados mediante una búsqueda en backward chaining.

O sea, todo antecedente trata de verificarse con alguna proposición de la base de hechos o la considerada como de inferencias, en dependencia del tipo de predicado asociado con el antecedente y definido en la

base de predicados. Si un antecedente no puede verificarse con un hecho, entonces trata de verificarse con algún consecuente de TD, la cual a su vez requiere verificar nuevos antecedentes, repitiéndose este proceso hasta la verificación total de los antecedentes. Si durante este proceso algún antecedente fallara, se analizarían los mismos mediante el proceso conocido como backtracking.

Además de estos mecanismos ampliamente conocidos, el modelo TDC incluye otros procesos específicos que complementan los anteriores: el case y el procesamiento conjunto de antecedentes o proceso alfa.

El case y el proceso alfa consideran las peculiaridades que del modelo TDC se han implementado, tales como:

- la existencia de una red jerárquica de categorías atómicas.
- la posibilidad de cuantificar universal o existencialmente dichas categorías.
- las inferencias lógicas y de herencia que se establecen por esos conceptos.
- el encapsulamiento y comunidad de antecedentes de las reglas contenidas en cada tabla.

En todo este proceso inferencial se implementarion funciones de propagación de los valores de certezas, las cuales se corresponden con el modelo teórico desarrollado por Hájek (6).

Además de los procesos inferenciales concebidas en el modelo TDC, este sistema permite implementar en Turbo Pascal procesos asociados con algunos predicados cuyos tipos fueron definidos como tal en la base de predicados. Estos procesos asociados son invocados al requerirse la verificación de un antecedente cuyo predicado sea de ese tipo.

De esta forma, el ingeniero de conocimientos pudiera diseñar bases que contemplen otros procesos específicos de su aplicación.

### Consideraciones finales.

Como fue planteado al inicio, el SETDC 1.0 es una primera aproximación al modelo TDC, en la cual ha sido implementado algunas de las características esenciales de ese modelo.

A pesar de la corta existencia de este sistema, se ha estudiado en un plano teórico el diseño de tres SE diferentes para:

- complementar procesos de asignación de recursos en una empresa de transportación.
- diagnosticar las posibles causas ante fallas en equipos y componentes de aviación.
- realizar la selección de rodamientos.

Este último SE se encuentra en la actualidad en las últimas fases de su implementación.

Como resultado de estos trabajos se ha verificado no solo la viabilidad del modelo y el sistema, sino que además ha corroborado las bondades esperadas.

Entre ellas debe resaltarse las ventajas que ofrecen las TD, tanto en su ejecución como en su carácter de instrumento de diseño y comunicación con los expertos, coadyuvando al completamiento y perfeccionamiento de las reglas y del conocimiento en general.

Tales trabajos también han permitido verificar las posibilidades de las categorías atómicas y su red taxonómica, aportando nuevos elementos los cuales se están analizando de una forma integral en sus dimensiones teórico-metodológicas.

Sin embargo, no obstante las posibilidades que ofrece esta primera versión, se ha comprobado la conveniencia de implementar nuevas

versiones del SETDC conteniendo otras características no contempladas del modelo TDC, tales como las categorías de objetos y las TD de entrada ampliada.

Los trabajos realizados y otros iniciados en la esfera del control de procesos tecnológicos han planteado nuevas exigencias al sistema como es la inclusión de procesos asociados no solo con antecedentes sino también con consecuentes, lo cual permitirá realizar SE con mayores posibilidades, requiriéndose quizás en un futuro próximo la normalización de tales procesos (7).

Muy asociado con lo anterior, se ha concebido la necesidad de desarrollar y perfeccionar, según las características del modelo, procesamientos forward-backward, lo cual potenciaría las posibilidades y tipos de aplicaciones a realizar en este sistema.

Todas estas características y otras se encuentran en un proceso de estudio para ser incluidas en futuras versiones del SETDC.

## Bibliografía.

- (1) Medina P., José E. "Sistemas de Tablas de Decisión". Ponencia de Informática 88. La Habana, 1988.
- (2) Medina P., José E. "El modelo de Tablas de Decisión Categorical: una representación del conocimiento en los sistemas expertos". Monografía. Rev. Ing. Industrial, Cuba (en proceso de edición).
- (3) Medina P., José E.; Castro D., Giselle y Alfaro R., Lisset. "SETDC 1.0 Manual de Usuario", 1990.
- (4) Marco, Tom de. "Structured Analysis and System Specification". Ed. Revolucionaria, La Habana, 1986.
- (5) Frenzel Jr. Louis E. "Crash Course in Artificial Intelligence and Expert Systems" S/A.
- (6) Hájek, P. "Combining functions for certainty degrees in consulting systems". Inter. Journal Man-Machine Studies. Academic Press Inc., London, 1985.
- (7) Díaz, Oscar; Jaime, Arturo. "Un gestor extensible de procedimientos asociados en el contexto de los sistemas orientados a objetos" IBERAMIA 92 La Habana, Cuba, Ed. Limusa, México, pp 59-74, 1992.