

Instituto Politécnico Nacional
UPIICSA ESCOM
México

SISTEMAS EVOLUTIVOS
BASADOS EN CONOCIMIENTO

Fernando Galindo Soria
21 de Enero de 1996

Sistemas Basados en Conocimiento

Sistema de Biblioteca Automatizada

Sistemas Expertos

Bases de Datos Inteligentes

Sistema Basado en Conocimiento

cuenta con diferentes mecanismos para representarlo y manejarlo, incluyendo:

- . directorios de datos.**
- . archivos de datos.**
- . bases de datos.**
- . bases de conocimiento.**
- . conjuntos de reglas de inferencia.**

Los medios de manejo de conocimiento se pueden integrar de diferentes formas

Sistemas de Base de Datos

Directorio de datos interrelacionado con la Base de Datos.

Sistemas Expertos o de Toma de Decisiones

integran al menos

Base de Conocimientos y un conjunto de reglas de inferencias.

Sistemas de Biblioteca Automatizada

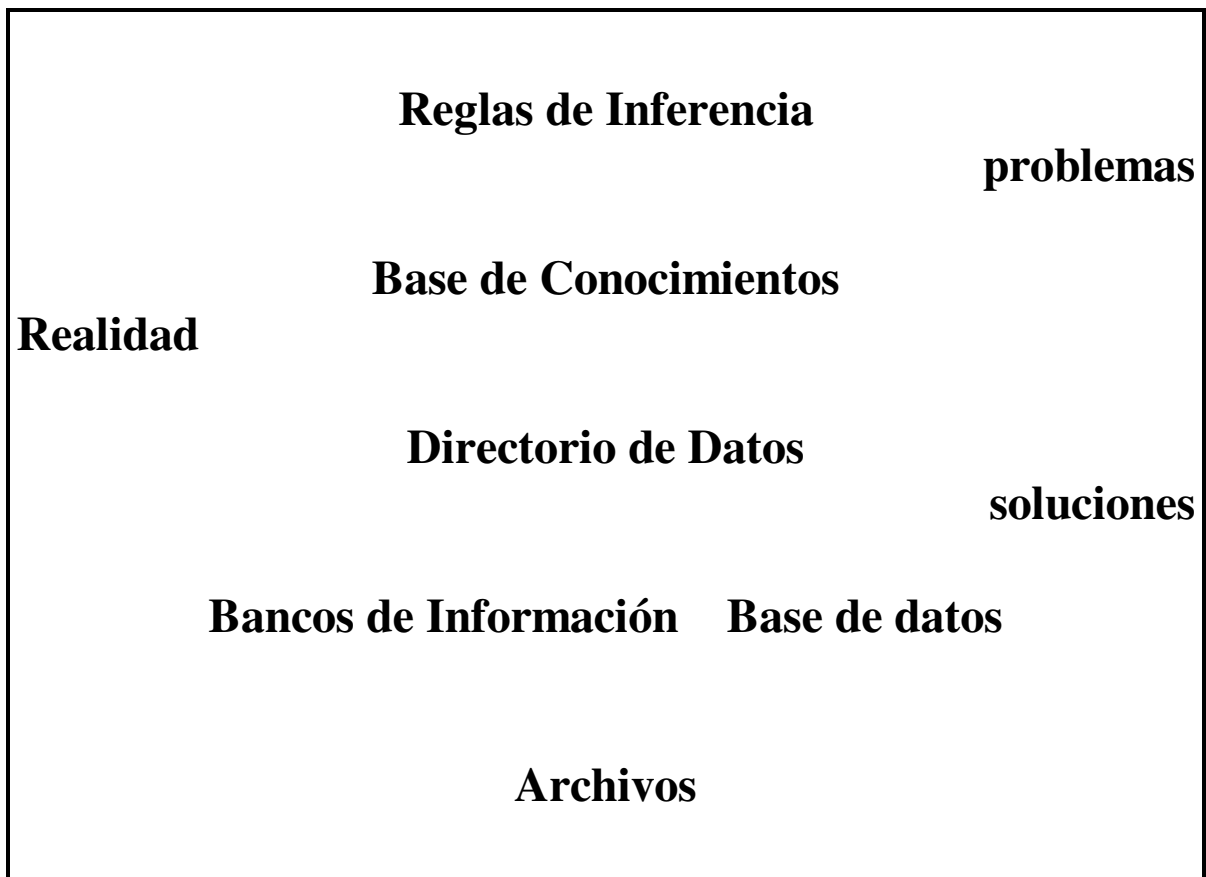
Formados al menos por

Bancos de información y directorios de datos

.

ARQUITECTURA DE SISTEMAS EVOLUTIVOS BASADOS EN CONOCIMIENTO

arquitectura por capas



**Mecanismos de Representación de Conocimiento relacionados
entre sí**

Mecanismos de percepción y Actuación

Problema muy grave

**Volumen masivo de datos
extremadamente complejo su tratamiento manual**

**El proceso de adquisición del conocimiento se debe
realizar en forma automatizada**

**Integración de Herramienta Evolutivas
orientadas a la adquisición y mantenimiento del sistema
basado en conocimiento**

**Sistema Evolutivo Basado en Conocimiento y bajo
una arquitectura por capas.**

GENERACIÓN AUTOMÁTICA DEL DIRECTORIO DE DATOS

Cada palabra clave en el directorio se necesita ubicar en los bancos de información para detectar en que párrafos se presenta, de tal forma que cuando un usuario pregunte por esa palabra el sistema le indique en que párrafos se encuentra.

Realizar lo anterior a mano es un proceso aburrido e impractico

Se necesitan herramientas que construyan los directorios en forma automatizada.

Durante los años 70's en El Colegio de México se inicio el Proyecto del Diccionario del Español en México

Se analizaron varios millones de documentos buscando las diferentes palabras y sus significado.

No se hubiera podido realizar manualmente

La cantidad de información a tratar era enorme

Se construyó, bajo la dirección del Ing. José Cen Zubieta

Un sistema de apoyo automatizado

Entraban los documentos a analizar

El sistema generaba para cada palabra la lista de los párrafos donde aparecía esa palabra

Herramienta General

Documentos almacenados en el banco de información

Primera capa del Sistema Evolutivo

**Toma los datos del banco y construye el directorio de
datos**

**Indicando para cada palabra clave la localización de los
párrafos donde aparece**

GENERACIÓN AUTOMÁTICA DE LA B. DE CONOCIMIENTOS

**Mecanismo que permite la construcción automática de la Base
de Conocimiento.**

La Base de Conocimiento

**Permite representar las interrelaciones entre las diferentes
palabra claves y su significado.**

Red Semántica

**Hechos o reglas de la forma
objeto relación objeto (o r o)**

Juan es hermano de Pedro

**de esta oración se obtienen directamente los componentes
de una red semántica.**

**Juan es un objeto
es hermano de es una relación
Pedro es otro objeto**

**En general los documentos no vienen divididos en párrafos de la forma
oro**

**Jesús Olivares
tesis de licenciatura "Sistemas Evolutivos para
Representación de Conocimiento" desarrollada en la
UPIICSA del IPN**

**Mecanismo evolutivo para construir
una Base de Conocimientos basada en una Red Semántica
Aumentada
(Que almacena datos, imagen y sonido).**

**Aplica una serie de técnicas de Inferencia Gramatical
basadas en la distribución lingüística**

**transforma una oración compuesta de múltiples
elementos en un conjunto de oraciones de la forma
oro
que son fácilmente representadas con una red semántica.**

En la oración:

Juan estudia en el poli, trabaja en el INEGI y vive en Lindavista.

se tienen varios hechos pero sin embargo no son directamente representables mediante una red semántica, se le aplicara la distribución lingüística

por facilidad primero representaremos los elementos de la oración mediante letras y números, de tal manera que se note directamente cuales son los objetos, cuales las relaciones y cuales los separadores:

Juan estudia en el poli , trabaja en el INEGI y vive en Lindavista

- **Juan** es un objeto o1
- **estudia en el** es una relación r1
- **poli** es un objeto o1
- **,** es un separador +
- **trabaja en el** es una relación r2
- **INEGI** es un objeto o2
- **y** es un separador +
- **vive en** es una relación r3
- **Lindavista** es un objeto o3

de donde la oración queda representada como:

$$o0 (r1o1 + r2o2 + r3o3)$$

si se toma la oración y se le aplica la operación de distribución queda:

$$o0 (r1o1 + r2o2 + r3o3) = o0r1o1 + o0r2o2 + o0r3o3$$

y si sustituimos el significado de los diferentes elementos tenemos que:

o0r1o1	Juan estudia en el poli
+	,
o0r2o2	Juan trabaja en el INEGI
+	y
o0r3o3	Juan vive en Lindavista

GENERACIÓN AUTOMÁTICA DE LAS REGLAS DE INFERENCIA

**Reglas de inferencia
componente básico de un Sistema Basado en
Conocimiento**

generación manual puede ser muy tediosa

**Generación automatizada de reglas de inferencia
trabajo desarrollado junto con Javier Ortiz en el
CENIDET de Cuernavaca**

En general una regla de inferencia es un regla de la forma:

Si A entonces B

A en comúnmente es un conjunto de condiciones o síntomas sobre un problema

B se puede ver como una acción o conjunto de acciones

en particular B es el diagnóstico y tratamiento asociado con los síntomas de A.

Para obtener las reglas de inferencia a partir de un documento:

Para cada párrafo donde se presente un conjunto de síntomas asociados con un diagnóstico y con un conjunto de tratamientos

se sustituye cada elemento por una representación interna donde se indica si un elemento es síntoma, diagnóstico o tratamiento e ignorando los elementos que no tengan significado

Al final queda una representación de la forma:

$s_1s_2s_3\dots s_n \quad d \quad t_1t_2\dots t_m$

esto se realiza para todos los párrafos del documento y al final se cuenta con un conjunto de reglas como:

$s_1s_2s_3\dots$	$.s_n$	d_1	t_1
$s_m\dots\dots$	$\dots s_x$	d_2	t_2
$s_y\dots\dots$	$\dots s_z$	d_z	t_z

Mediante la Factorización Lingüística se pueden encontrar las reglas de inferencia del sistema basado en conocimiento

Factorización lingüística

si factorizamos o0 de la oración

$$\underline{o0}r1o1 + \underline{o0}r2o2 + \underline{o0}r3o3$$

nos queda

$$\underline{o0} (r1o1 + r2o2 + r3o3)$$

Se toman las reglas encontradas anteriormente y se factorizan los elementos comunes

los síntomas comunes a varios diagnósticos quedan juntos y los síntomas particulares de un diagnóstico quedan asociados con él.

CONCLUSIÓN

**Partimos de que se cuenta con
un banco de información al que se le aplican técnicas
evolutivas para construir
el Directorio de Datos,
la Base de Conocimiento
y el conjunto de Reglas de Inferencia.**

**Lo maravilloso de este enfoque es que
a partir de la misma fuente de información se pueden
encontrar diferentes niveles de representación de
conocimiento y**

**lo que se esta realizando es un diálogo con el interior del
Banco de Información**

**para encontrar reglas y conocimiento que se encuentra
inmerso en las fuentes de información y que por lo común
se pierden.**

Sólo es necesario construir herramientas capaces de realizar procesos de diálogo con los almacenes de información

y estos mecanismos son capaces de encontrar una gran cantidad de información inmersa

que hasta el momento no se aprovecha por no contar con las herramientas adecuadas

Estamos sentados sobre una mina de información y no sabemos aprovecharla o siquiera que existe.