

UN ROBOT CAPAZ DE DIALOGAR EN CASTELLANO (2º PARTE)

Por M.F.Verdejo, Analista del CCUM.

1-Introduccion

Nuestro trabajo se plantea el estudio de un proceso completo de "comprensión", en un modelo formado por objetos y un robot que puede realizar determinadas acciones sobre ellos.

En el diálogo que se entabla en castellano, el robot recibe órdenes que debe ejecutar y responde a las preguntas que se le formulen relativas a su mundo.

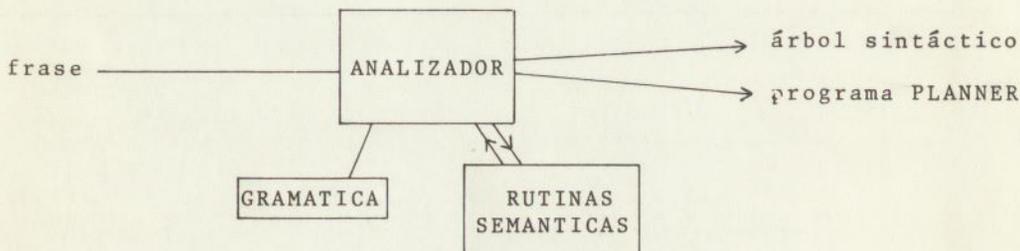
El lenguaje computacional que hemos elegido para representar la información en el ordenador, está inspirado en PLANNER, el proceso de resolución (responder a las preguntas, ejecutar las órdenes) consta de :

- 1) Identificación del objeto (u objetos) referido.
- 2) Una serie estructurada de acciones para llevar a cabo la tarea pedida, expresadas mediante un conjunto básico de operaciones que en el caso del robot son MOVER, LIBERAR, COGER, ect.

Con este planteamiento, el problema que queda por resolver, es la traducción de las frases producidas en castellano a su correspondiente programa PLANNER.

El proceso se realiza mediante un analizador sintáctico-semántico. El analizador sintáctico guía el proceso de reconocimiento de la frase estructurándola de acuerdo a las reglas de la gramática que se ha definido. El analizador semántico colabora en el análisis y al mismo tiempo construye el programa PLANNER que expresa el significado de la frase.

El esquema de esta etapa es el siguiente :



2.- El analizador semántico

El analizador semántico construye el programa PLANNER, representación interna de la frase, a partir del significado de las palabras que la componen.

Hay dos formas básicas de definición semántica de las palabras, una que llamamos standard, consiste en la definición PLANNER correspondiente, por ejemplo, la palabra pirámide, tiene una entrada en el diccionario semántico, que contiene la expresión :

PIRAMIDE (\$PIRAMIDE) (\$ES - \$PIRAMIDE)

En el primer paréntesis, aparece el marcador semántico que expresa las propiedades del objeto, este marcador es un elemento de una jerarquía de atributos específica del modelo, así por ejemplo \$PIRAMIDE, implica \$MANEJABLE y \$OBJETO FISICO.

El segundo paréntesis es la expresión en PLANNER.

Los verbos, por ejemplo "coger" tienen una entrada del tipo:

COGER (\$ANIMADO \$MANEJABLE) (\$COGER - -)

En el primer paréntesis, se indica las propiedades de los objetos que participan en la acción. El que la realiza debe ser \$ANIMADO, sobre un \$MANEJABLE. En el segundo paréntesis aparece la expresión PLANNER, las barras se sustituirán en el proceso de análisis por los elementos adecuados, para ello, se comprueba que los marcadores semánticos sean compatibles.

Por ejemplo, dada la frase : Coge la pirámide verde.

Cuando se analice el grupo nominal "la pirámide verde", el programa semántico que trata esta unidad sintáctica, comienza la construcción de una estructura con los siguientes campos :

En el primero, la expresión PLANNER, que se crea a partir de la definición del nombre, así después de haber analizado "la pirámide", tenemos :

THFIN (\$ES X \$PIRAMIDE)

En el segundo campo, la lista de marcadores semánticos extraídos de la estructura jerárquica a partir de \$PIRAMIDE, que son :
\$MANEJABLE \$OBJETO FISICO

El tercero contiene información sintáctica, en este caso, se trata de un GNOM, con las características singular y definido.

El cuarto contiene la variable elegida para representar el objeto, en el ejemplo, X.

Al analizar la palabra siguiente "verde", el especialista semántico comprueba que es una propiedad que puede tener el objeto comparando los marcadores semánticos; La definición del adjetivo es

VERDE (\$OBJETO FISICO) (\$COLOR _ \$VERDE)

propiedad que se aplica a un objeto físico, este marcador se encuentra en la estructura que estamos construyendo, luego podemos añadirla, y el programa será al final del análisis del Grupo nominal :

THPROG (X)

THFIN (\$ES X \$PIRAMIDE)

THFIN (\$COLOR X \$VERDE)

Hay otras palabras, como es el caso del verbo ser, o los pronombres, que pueden referirse a objetos, hechos sucedidos, personas y que por lo tanto no vienen definidas por una expresión, sino que su significado lo determina el contexto en el que aparecen. Así por ejemplo, en la frase:

"Cogelo", el pronombre se refiere al objeto del que estamos hablando. Mientras que :

¿Cómo lo has hecho? se refiere a una acción realizada.

Para estas palabras, existen programas especialistas, que determinan la representación interna, estudiando el contexto; En el primer caso, los marcadores semánticos del verbo coger, indicarian acción sobre un objeto físico, consultando la "memoria inmediata" y aplicando los heurísticos disponibles, se resolvería la referencia en favor del objeto correspondiente.

En el segundo caso, los marcadores del verbo hacer, indicarian \$ACCION, y se determinaría la referencia consultando la memoria de acciones realizadas.

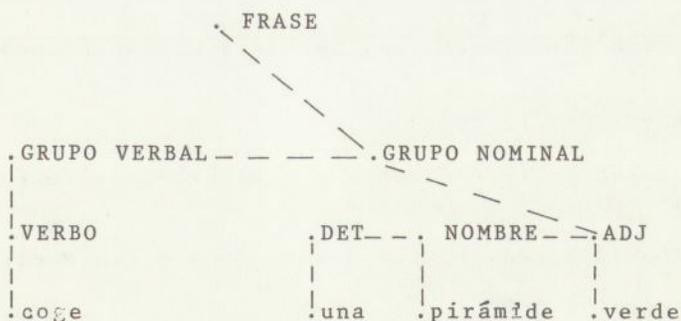
El desarrollo del proceso, se hace bajo el control del analizador sintáctico que estructura la frase en unidades, desde las que se llaman a los programas semánticos que construyen las expresiones PLANNER correspondientes a cada una de ellas.

Estas unidades, están a su vez, elegidas por criterios semánticos. En una frase, nos interesa determinar los objetos, que gramaticalmente serán grupos nominales. Estudiar sus propiedades y relaciones que vendrán expresadas mediante grupos preposicionales adjetivales y verbales.

Así por ejemplo, la frase imperativa :

(Coge una pirámide verde.

Tiene una estructura sintáctica, que puede representarse por el árbol :



En el proceso de análisis del Grupo Nominal (GNOM), se construye la lista de expresiones PLANNER correspondiente a las palabras que la componen. Al llegar a fin de grupo, el especialista semántico estudia las características sintácticas indicadas en el tercer campo de la estructura para determinar la cabecera PLANNER del programa, así por ejemplo, para el grupo "una pirámide verde", el operador es un THPROG (X).

Mientras que para "todas las pirámides verdes", el operador sería un THBUSCAR TODO (X).

Igualmente el grupo verbal crearía la expresión PLANNER correspondiente al verbo, en el ejemplo:

THFIN (\$COGER -) THUTILIZAR (COGER)

El especialista semántico de FRASE, es el que reorganiza y construye el programa definitivo. Para ello tiene en cuenta la información contenida en cada una de las estructuras semánticas, la estructura temporal de la frase, y una serie de heurísticos de ordenación. En el ejemplo, el programa sería :

THPROG (X)

THFIN (\$ES X \$PIRAMIDE)
 THFIN (\$COLOR X \$VERDE)
 THFIN (\$COGER X)
 THUTILIZAR (COGER)

Las frases interrogativas, formadas con un pronombre, como por ejemplo :

¿Qué hay sobre la mesa?

Se representan internamente por el programa correspondiente a la frase :

Buscar todos los objetos que se encuentren sobre la mesa.

Las preguntas del tipo :

¿Hay una pirámide más grande?

Que tienen como respuesta si ó no, se representan por la frase equivalente :

Buscar (el grupo nominal), en este caso:

Buscar una pirámide más grande.

Hay frases en las que la determinación del GNOM principal, no es evidente, como por ejemplo :

¿Hay algún bloque que soporte otro que soporte algo?

Los criterios que se siguen para la determinación del Grupo Nominal principal son dos :

- 1° Elegir un GNOM que sea indefinido.
- 2° Elegir un GNOM que no esté imbricado en otro grupo.

Estas dos reglas de decisión permiten determinar el Grupo respecto al cuál vamos a realizar la transformación.

3.- El tiempo

En la memoria del robot se guardan las acciones que éste ejecuta, con los pasos intermedios más importantes. Los cambios de posición de los objetos, con indicación del momento en que se llevaron a cabo. Esta información se utiliza para contestar a las preguntas sobre los hechos pasados, el momento en que se realizaron reconstrucción de escenas, ect.

Ejemplos de este tipo son las preguntas :

- ¿Has cogido algún cubo?
- ¿Cuántos objetos habia en la caja?
- ¿Cuántos bloques habia debajo de la pirámide que has movido?
- ¿Cuándo lo has movido?
- ¿Por qué lo has hecho?

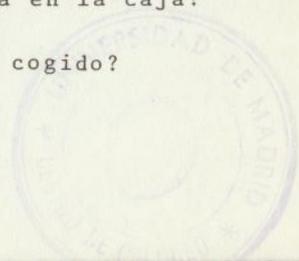
Durante el proceso de análisis, para cada frase, se crea una estructura temporal, en la que se anotan todas las características relativas al tiempo, implícitas en el grupo verbal, en los adverbios, ect.

El especialista semántico correspondiente a la unidad sintáctica FRASE, al realizar la organización de los programas construidos por las unidades de nivel inferior, consulta esta información temporal, modificando convenientemente las cabeceras PLANNER.

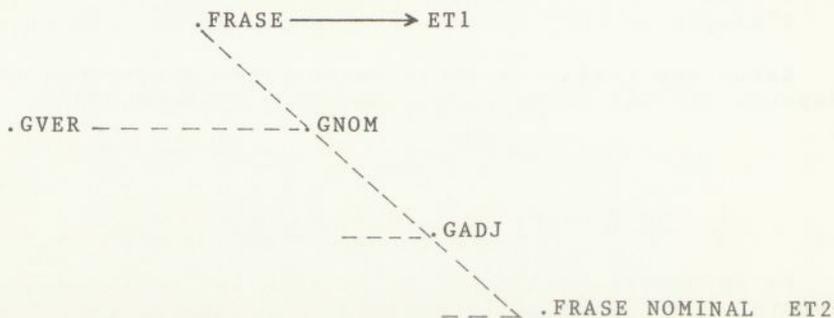
Cuando el espacio de búsqueda de soluciones es la MEMORIA, el operador PLANNER es un THMEMO, que funciona igual que el THFIN sobre la escena actual.

Así por ejemplo, las frases :

- ¿Hay algún bloque más pequeño que el que está en la caja?
- ¿Hay algún bloque más pequeño que el que has cogido?



Tienen la estructura sintáctica siguiente:



ET- estructura temporal

Mientras en la primera frase ET2 marca tiempo presente, en la segunda es el pasado. Lo que da lugar a dos programas diferentes que son los siguientes :

THPROG (X Y Z)

```

(THFIN ($ES X $BLOQUE)
THFIN ($ES Y $CAJA)
THFIN ($CONTIENE Y X)
THBUSCAR (1) (Z)
  ( THFIN ($ES Z $BLOQUE)
    THFIN ($MAS $TAMAÑO X Z)
    THUTILIZAR (MEDIDA)))
  
```

THPROG (X Y Z)

```

(THFIN ($ES X $BLOQUE)
THMEMOR($COGER X)
THBUSCAR (1) (Z)
  (THFIN ($ES )Z $BLOQUE)
    THFIN ($MAS $TAMAÑO X Z)
    THUTILIZAR (MEDIDA)))
  
```

4.- El contexto

Las palabras que componen el lexico del modelo, aparecen definidas o bien en forma standard en el diccionario semántico o mediante programas que determinan su significado estudiando el contexto en el que aparecen.

Este es el caso de los verbos ser, estar y de la forma impersonal "hay", ejemplos de frases son los siguientes :

¿Cuántos cubos hay sobre la mesa?

El bloque que está a la derecha de la pirámide

¿De qué color es?

El bloque azul es mio

¿Cuántos bloques han sido desplazados?

En nuestro modelo, el verbo "estar" aparece indicando la posición o situación física de un objeto. Al igual que la forma "hay" es el grupo preposicional el que marca la relación: "sobre", "a la derecha", "encima de"...

El verbo ser aparece como auxiliar, o indicando una propiedad, que puede ser el color, la forma, la pertenencia, ect. Su significado se representa por la expresión interna de esa propiedad.

En el caso de los pronombres, de las frases y grupos incompletos, el estudio del contexto es fundamental para determinar las referencias, que pueden ser a :

- Objetos concretos
- Descripciones de objetos
- Hechos sucedidos

Ejemplos de frases de este tipo son las que se producen en el diálogo siguiente :

- 1- Pregunta - ¿Qué hay en la caja?
- 2- Respuesta- La pirámide azul.
- 3- Orden - Ponla sobre el bloque verde
- 4- Respuesta- De acuerdo.
- 5- Pregunta - ¿Cómo lo has hecho?
- 6- Respuesta- Desplazando primero el cubo que estaba sobre el bloque verde.
- 7- Pregunta - ¿Hay alguna más grande?
- 8- Respuesta- Si, la pirámide roja.
- 9- Pregunta - ¿Dónde está?
- 10- Respuesta- Sobre la mesa.

El pronombre "qué", en 1, se refiere a objetos movibles, el pronombre "la" en 3, se refiere a la pirámide azul, el pronombre "lo" en la frase 5, se refiere a la acción realizada.

En la frase 7, hay que resolver dos referencias, una respecto a "alguna" que significa "alguna pirámide", otra respecto al grupo adjetival incompleto "más grande que la pirámide azul".

La frase 9, también es incompleta. La pregunta se refiere a la pirámide roja.

Para resolver estas referencias, el analizador construye y utiliza una "memoria inmediata" en la que va guardando la definición semántica y la identificación de los objetos de los que se habla.

Por medio de las características sintácticas y los marcadores semánticos se determina en cada caso la identificación del objeto.

5.- La gramática

La finalidad del análisis sintáctico es organizar la frase en unidades que tienen una función semántica definida.

A un primer nivel, tenemos CLASES de palabras, que se combinan en grupos sintácticos, que a su vez forman frases.

La gramática que utilizamos describe cada una de estas unidades : palabra, grupo, frase, especificando las formas posibles y su función, mediante una organización jerarquizada de "Características" sintácticas.

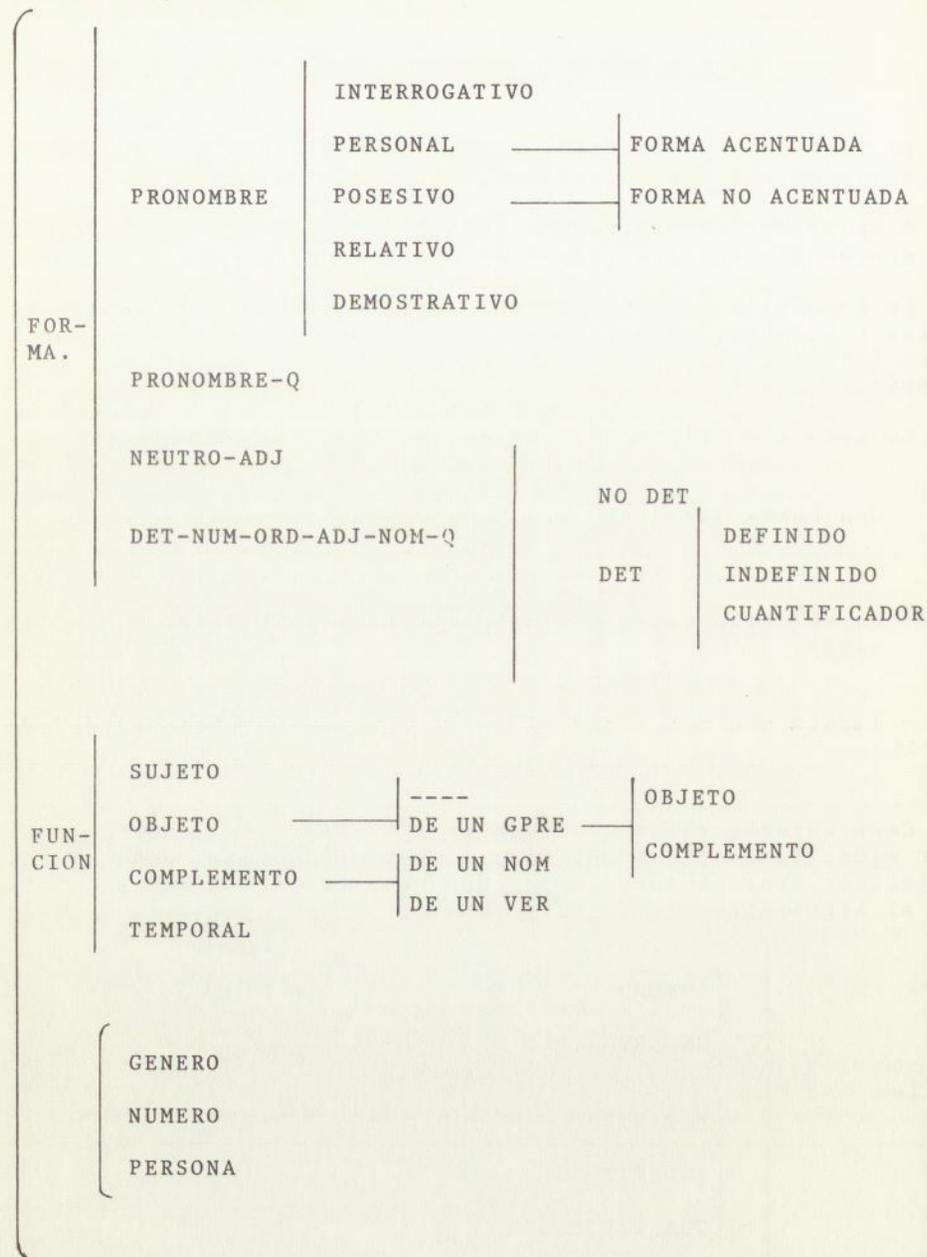
La notación utilizada, emplea los símbolos siguientes :

- Una barra vertical para indicar elección única.
- Un corchete para indicar elecciones múltiples, no exclusivas.
- Indica que dos o más elementos comparten características.

Cada unidad, tiene asociada un SISTEMA de características. Por ejemplo, la clase DET formada por artículos y pronombres adjetivos, aparece como componente del Grupo Nominal, su sistema es el siguiente :

{	GENERO
	NUMERO
	DEFINIDO
	INDEFINIDO
	CUANTIFICADOR

El del Grupo Nominal :



6.- El analizador sintáctico

El analizador es una red recursiva de autómatas finitos, que realiza un análisis TOP-DOWN, guiado por las características sintácticas que aparecen en los sistemas de definición de la gramática, en continua interacción con las rutinas semánticas.

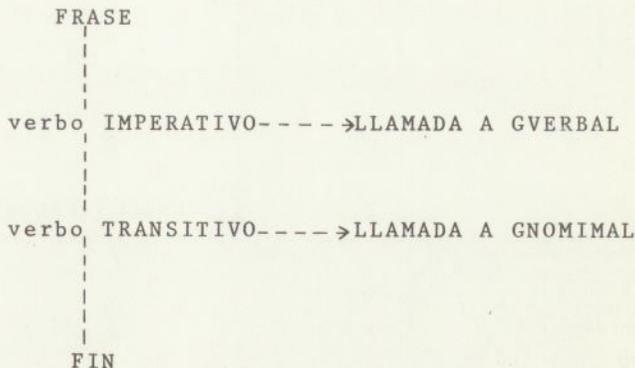
En el transcurso del análisis se reconoce la frase, produciendo una arborescencia sintáctica y se construye una estructura semántica, que contiene en uno de sus campos, el programa PLANINER, representación interna de la frase.

El proceso, comienza con la activación del autómata FRASE, que será el que llame (en proceso recursivo), a los demás autómatas.

Por ejemplo, el análisis de la frase :

"Coge la caja que está sobre la mesa"

Empieza con la activación de FRASE, el diagrama es el siguiente:



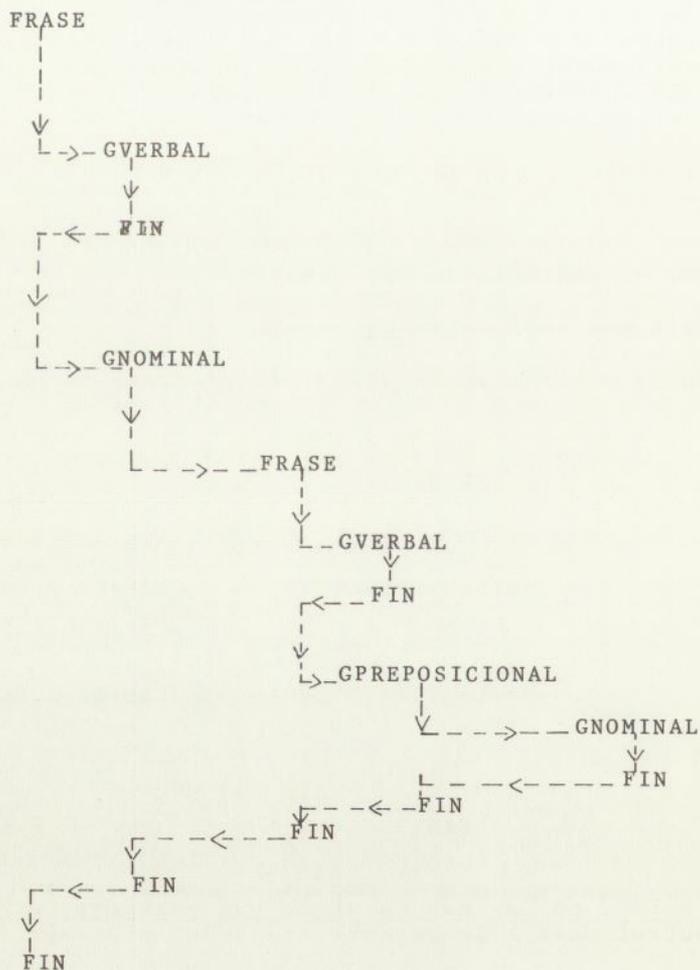
Al llegar a FIN, no hay más palabras que analizar, el proceso termina, el control pasa a la segunda etapa del sistema.

En la llamada al GNOMINAL, las características DET DEFINIDO, hacen escoger la forma

DET-NUM-ADJ- NOM- Q

El análisis de la palabra "que", con las características de relativo, provocan la llamada al autómata FRASE, con indicación de vuelta al Grupo Nominal, finalizado el control de FRASE.

El esquema completo de la ejecución par este caso, es el siguiente :



El analizador está implementado en un lenguaje que hemos diseñado para facilitar el tratamiento de los datos y las operaciones que se llevan a cabo sobre ellos, de forma que las modificaciones sean sencillas. Las instrucciones de este lenguaje son fundamentalmente de dos tipos, unas trabajan sobre cadenas de símbolos (el input), las otras sobre la estructura arborescente (el output).

Cada instrucción se atiende al siguiente formato :

Primer Campo - (opcional), etiqueta

Segundo Campo - Código de la operación

V - Verificar una característica del símbolo (palabra) que se procesa de la cadena de entrada.

P - Creación de un nodo en el árbol sintáctico.

LEC- Llamada al analizador morfológico.

A - Añadir características a un nodo del árbol.

B - Buscar un nodo con una característica

SMT- Llamada al analizador semántico.

Tercer Campo - Dependiendo del tipo de instrucción será o una cadena de letras, o una cadena de letras y asteriscos.

Ejemplo : Verificar si la palabra que hemos leído es un pronombre

V PRO

Añadir al padre del nodo, la característica SINGULAR.

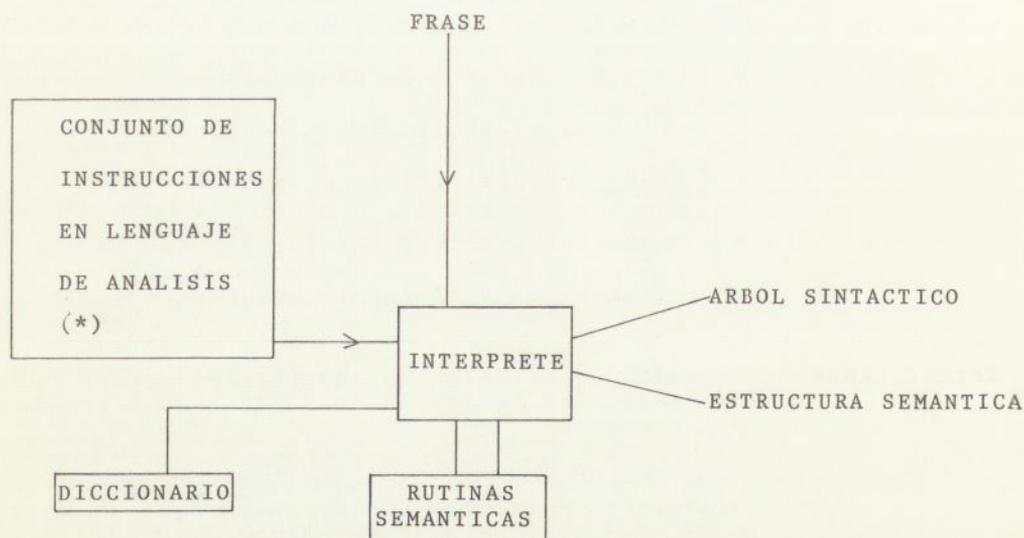
A P*N SIN

El campo cuarto (opcional), indica la característica que hay que añadir, buscar o verificar. Los campos quinto y sexto indican la transferencia de control en caso de que la instrucción se realice con éxito o fracaso.

Para más detalle, ver referencia (8).

Los autómatas se transcriben fácilmente por este lenguaje; junto con la frase a analizar, constituyen el INPUT del intérprete .

El esquema es el siguiente :



(*) Constituyen la red de autómatas que reconoce las frases del subconjunto del castellano que tratamos.

7.- Conclusión

Para implementar un sistema de diálogo en lenguaje natural, tan simple como el modelo del robot, es necesario plantearse y resolver procesos tales como el reconocimiento del lenguaje, la forma más adecuada de representar el significado, la manera de organizar el conocimiento en la memoria (qué información se retiene, qué información se olvida), el tipo de inferencias imprescindibles para generar una respuesta que no se encuentre de forma explícita en la representación interna del modelo.

La complejidad en el estudio de los procesos cognoscitivos, reside en la cantidad de información que se trata (su selección y organización) , y en la interacción de los diferentes niveles (sintaxis-semántica-pragmática) que intervienen.

La elección de una aplicación simple se base en dos razones. La primera, es que en un modelo del tipo del robot, la información contextual es limitada, las ambigüedades que se presentan son fáciles de resolver. Esto permite clarificar la conexión e importancia de cada una de las fases del proceso de "comprensión". El papel del análisis sintáctico, qué se entiende por semántica, por qué una representación es adecuada, cuáles son sus limitaciones, qué cantidad de información previa es necesaria, cuáles son los mecanismos más útiles para manejarla.

La segunda, es que siguiendo un criterio de generalidad, se pueden desarrollar instrumentos potentes de manejo de datos que sean útiles a otras aplicaciones.

Bajo este punto de vista, los resultados obtenidos han sido satisfactorios. De la comparación de nuestro estudio con el que realizó Winograd (10), podemos afirmar que el método seguido no depende de la lengua que se utilice, tanto del inglés como del castellano, las frases se pueden traducir al lenguaje PLANNER.

Evidentemente, el estudio de la semántica y de la gramática es diferente en cada caso. En castellano, el analizador morfológico es más complejo, pero esto favorece al analizador sintáctico, ya que las palabras al presentar más características (género, número, persona, etc.) proveen más información.

El orden de composición de los elementos que forman la frase, es más libre que en inglés y por tanto el analizador tiene que ser más flexible en cuanto a las predicciones de las palabras que pueden seguir a la que se está procesando.

Las rutinas semánticas son más elaboradas. En castellano, el estudio de referencias (debido a la nominalizaciones de los adjetivos, al amplio uso de los pronombres) tiene gran importancia en el transcurso del diálogo.

El modelo puede representarse por una cantidad limitada de información. La calidad de los resultados responde a la adecuación del lenguaje PLANNER, para expresar sin ambigüedad, por un número reducido de atributos, los objetos del modelo. El conocimiento necesario para manejarlos se representa por una composición de acciones elementales.

La memoria registra las acciones y pasos intermedios importantes, con referencia al tiempo en que se ejecutaron.

Al ampliar el campo, y escoger una aplicación del mundo real, las limitaciones más importantes provienen de esta organización simplista de la memoria.

Del estudio de los mecanismos cognoscitivos en el hombre, podemos observar que el conocimiento previo que poseemos juega un papel muy importante en los procesos de comprensión. Lo que ya sabemos, permite esperar o predecir lo que ocurrirá en situaciones normales, de forma que cuando se establece una comunicación, el reconocimiento (perceptivo o mediante el lenguaje) es un proceso altamente predictivo, basado en la experiencia que hemos adquirido.

Además, podemos resaltar, que este conocimiento previo, no es únicamente un conjunto de hechos "verdaderos", hay información probablemente cierta o con diversos grados de credibilidad.

En los últimos tres años, los estudios en Inteligencia Artificial, se han centrado en la investigación de estructuras adecuadas de organización de la memoria.

Los trabajos en curso, las publicaciones más recientes (por el momento a un nivel exclusivamente especulativo), desarrollan diversas teorías, casi todas basadas en las ideas de los "frames" de Minsky (5).

Es de esperar, con fundamentado optimismo, que en un periodo corto de tiempo, estas ideas se plasmen en la realización de técnicas que permitan implementarlas. Esto constituirá un empuje importante en el camino hacia la construcción de sistemas que tengan un comportamiento "inteligente".

Bibliografia

- 1- Chomsky- Aspects of the theory of Syntax
MIT Press 1965
- 2- Green - Baseball, an automatic question-answerer
Proc. Western Joint Computer conf.
- 3- Halliday Functional diversity in language as seen from a con-
sideration of modality and mood in English.
Foundations of language, 6, pg 322-361
- 4- Hewitt - PLANNER: a language for proving theorems in robots
Proceedings of the International Joint Conf. on
Artificial Intelligence. 1969. Pg 295-301.
- 5- Minsky - A framework for representing knowledge.
The Psychology of computation. N. York .Mc Graw Hill 1975
- 6- Robinson A machine oriented logic based on the Resolution
Principle.
JACM vol 12, n°1, 1965
- 7- Thorne - The Syntactic Analysis of English by machine
Machine Intelligence 3, N. York 1968
- 8- Verdejo - Une étude du langage naturel,
Thèse 3° cycle, Univer. Paris VI, 1975.
- 9- Weizenbaum ELIZA: a computer program for the study of natural
language communication between man and machine.
CACM Vol 9, pg 36-45, 1966
- 10- Winograd Understanding natural language
Academic Press N. York 1972
- 11- Woods - Transition Network Grammars for Natural Language
analysis.
CACM Vol 13, pg 591-606, 1970