

UN ROBOT CAPAZ DE DIALOGAR EN CASTELLANO (\*)

Por Maria Felisa Verdejo, Analista del Centro de Cálculo de la Universidad Complutense.

El presente trabajo plantea el problema de la comunicación con el ordenador en lenguaje natural.

El sistema realizado acepta órdenes y preguntas, formuladas en castellano, relativas al modelo implementado : un robot que demuestra su "capacidad de comprensión" efectuando las acciones pedidas o respondiendo adecuadamente a las cuestiones planteadas:

Hemos dividido la exposición en dos partes, la primera consta de :

- 1) Una introducción general
- 2) Presentación del modelo
- 3) Lenguaje utilizado para el tratamiento automático de la información , el PLANNER.

En la segunda abordaremos el método de análisis que permite obtener para cada frase producida en el diálogo, su correspondiente expresión en PLANNER.

---

(\*) La segunda parte de este trabajo aparecerá en el próximo número de este Boletín.

## 1- INTRODUCCION

Los primeros ordenadores se concibieron como instrumentos al servicio del cálculo científico. Eran máquinas que podían ejecutar con rapidez las operaciones indicadas por un conjunto de instrucciones (programa), sobre unas clases de operandos de tipo numérico.

La memoria permitía guardar los datos sobre los que se iba a efectuar el cálculo y además las instrucciones (las operaciones a realizar sobre ellos), la unidad de control ejecutaba las órdenes depositadas previamente en la memoria, de aquí que el proceso fuera automático.

Tanto los datos como el programa se representaban en una codificación basada sobre el alfabeto binario.

La máquina creada tenía unas posibilidades que ampliaban el campo de trabajo inicial, podía realizar cualquier tipo de proceso expresable mediante un repertorio de instrucciones y que trabajara sobre información representable en el ordenador.

La traducción automática fué una de las aplicaciones que más interés despertó, resumiendo brevemente su historia, podemos distinguir diferentes etapas : Los primeros intentos datan de 1950, numerosos investigadores abordaron el tema pensando que se obtendrían resultados aceptables que permitirían una operatividad inmediata.

La orientación de estos primeros trabajos, se centraba en la idea de explotar la capacidad de memoria y rapidez de la máquina para manejar grandes diccionarios en un proceso de traducción palabra por palabra.

Los resultados obtenidos fueron decepcionantes, una palabra es un elemento de una estructura ( grupo, frase o discurso...), y su significado está en relación con otros elementos, considerada en si daba lugar a traducciones aberrantes.

Algunos investigadores intentaron añadir información sobre ciertas palabras por medio de tablas, para intentar mejorar la calidad del resultado.

El final de esta primera época viene marcado por un desinterés general por el tema, que perduró durante un cierto tiempo.

La segunda etapa se abre con una perspectiva de investigación más amplia. De una aplicación en concreto ( la traducción automática ), ahora el estudio se centra en el lenguaje natural, su análisis y representación en el ordenador para la construcción de sistemas de comunicación hombre-máquina.

Se perfilan dos líneas de trabajo, la primera se orienta en el campo de la lingüística intentando aplicar las nuevas teorías de Chomsky sobre formalización de la sintaxis y definición de las propiedades matemáticas de las gramáticas. A partir de un subconjunto del lenguaje, diferenciando los niveles morfológicos, sintáctico y semántico se desarrollan ampliamente los dos primeros.

Así en constante evolución aparecen analizadores automáticos que reconocían las frases generadas por gramáticas regulares ( los autómatas de estados finitos ), las gramáticas de contexto libre, y las transformacionales ( trabajos de Thorne ), en donde el núcleo era una gramática de contexto libre.

Con el conjunto de reglas transformacionales se hacía un análisis de superficie para obtener la estructura profunda de la frase.

Siguiendo esta orientación aparecen también las redes aumentadas de transición, de Woods.

La otra línea, que es la que a nosotros nos interesa, se desarrolla en el contexto de lo que se ha llamado Inteligencia Artificial.

Partiendo de un modelo se intenta crear un sistema en el que la máquina se comporte como un interlocutor "inteligente". Esto implica la comprensión del lenguaje en el que el diálogo se produzca y una cierta capacidad de razonamiento para producir las respuestas necesarias.

Citaremos los trabajos más importantes, señalando los rasgos que caracterizan su evolución.

Entre los primeros que aparecen se encuentra BASEBALL de Green (1961), un sistema que respondía a las preguntas que se hicieran en inglés sobre los tanteos de los partidos producidos en este juego.

Las frases que se empleaban tenían que respetar unos formatos preestablecidos. Los resultados les dieron un cierto éxito, pero su limitación lingüística, y sobretodo lógica, era evidente. El sistema no estaba dotado de un procedimiento general capaz de hacer deducciones independientemente de la base de datos sobre la que se realizaba la aplicación.

La idea que surgió para solucionar este problema fué la de utilizar un lenguaje semiformalizado para representar y manejar la información en la máquina. En esta línea aparecen los trabajos que podemos clasificar de "lógica limitada", sistemas como ELIZA de Weizembaum (1966), que no solo acepta información sino que es capaz de responder a preguntas que impliquen relaciones lógicas elementales.

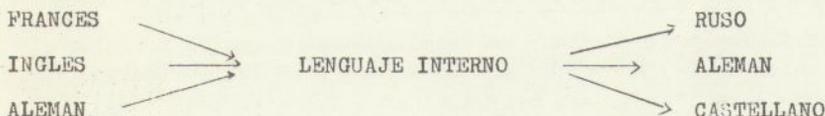
En esta misma época, los trabajos de Robinson (principio de resolución, 1965) constituyeron un avance fundamental en la construcción de demostradores automáticos de teoremas. Se trata de una regla de inferencia para encontrar pruebas en la lógica de primer orden. A partir de cualquier base de axiomas, se puede demostrar una fórmula aplicando un procedimiento general que permite encontrar la solución, si ésta existe, en un tiempo finito.

La elección de un lenguaje formal para representar internamente las información permite aplicar los métodos de demostración automática. Así por ejemplo, una pregunta se convierte en una fórmula a validar utilizando la base de axiomas que describe el modelo sobre el que se experimenta.

En el caso concreto de la traducción automática, este planteamiento nos lleva a definir un proceso con dos pasos, en el primero se traduciría del lenguaje natural fuente al lenguaje escogido como representación interna, en el segundo, de éste al lenguaje natural objeto.

LENGUAJE FUENTE  $\longrightarrow$  LENGUAJE REPRESENTACION  $\longrightarrow$  LENGUAJE OBJETO  
INTERNA

Con este tratamiento se evita tener un traductor bidireccional para cada par de lenguajes naturales, lo que se desarrollaría, sería uno lenguaje fuente  $\rightarrow$  l. interno, otro l. interno  $\rightarrow$  lenguaje fuente.



El primer paso sería un proceso de análisis (extraer el significado de una frase del lenguaje fuente), el segundo, de generación: a partir del "significado" representado por el lenguaje interno, producir la frase en la lengua correspondiente.

## 2-PRESENTACION DEL MODELO IMPLEMENTADO

Para abordar el estudio de un proceso completo de comunicacion, es necesario elegir una aplicaci3n que limite las infinitas posibilidades de un di3logo en castellano, de forma que aunque el vocabulario quede restringido, las caracteristicas esenciales del lenguaje permanezcan para que el resultado presente un cierto inter3s.

Nosotros hemos escogido el modelo siguiente :

Suponemos un robot ( el problema de su realizaci3n ffsica no nos interesa ), situado en una habitaci3n, en donde se encuentran diversos objetos, unos se pueden mover como los bloques y las pir3mides, otros no, la mesa , la caja. Sobre los primeros el robot realizar3 una serie de acciones.

Estos objetos se distinguen por el nombre que le damos a cada uno de ellos o por una descripci3n que los identifique, asf por ejemplo podemos referirnos a un objeto como :

" El objeto N3mero 1 "

o bien " El bloque verde que est3 sobre la mesa "

El di3logo que vamos a llevar a cabo ser3 en castellano, y en 3l, daremos 3rdenes al robot que tendr3 que comprender y ejecutar, o le haremos preguntas sobre la situaci3n actual de los objetos, sus caracteristicas, lo que hace, ect.

Basicamente ser3 un conjunto de frases :

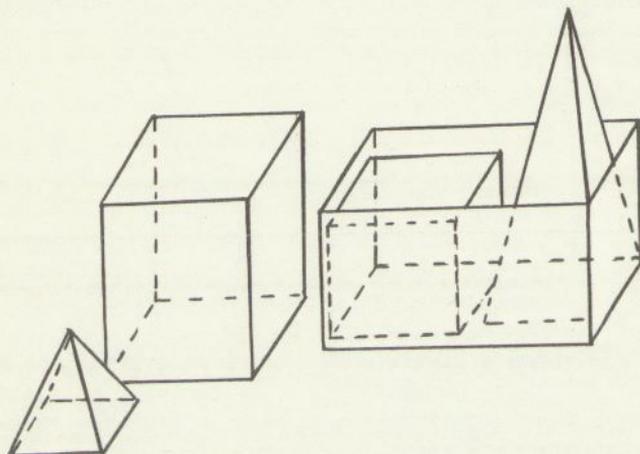
PREGUNTA - RESPUESTA

ORDEN - ACCION + RESPUESTA

Lo que implica que no tratamos las frases declarativas que serian 3tiles en un contexto de aprendizaje que nosotros no incluimos en nuestro modelo.

El tiempo viene representado por una variable T que toma el valor cero al comienzo del di3logo y se incrementa en una unidad cada vez que el robot ejecuta un movimiento.

Sobre una situación inicial como la de la figura, un ejemplo de diálogo que se produce es el siguiente :



Pregunta - ¿Qué hay en la caja ?

El robot responde : El cubo rojo, la pirámide azul

Orden - Coge el cubo rojo

Es una acción que efectúa desplazando el brazo a la posición en donde está el cubo. Como respuesta da las nuevas coordenadas del objeto.

Orden - Ponlo sobre la mesa

En esta orden se hace referencia a un objeto de la frase anterior, el robot determina el objeto y lleva a cabo el desplazamiento.

Orden - Pon la pirámide verde encima

En este caso el objeto está definido, el lugar donde dejarlo es una referencia a la frase anterior.

Orden - Pon la pirámide azul sobre el bloque más grande

Primero busca "el bloque más grande", después realiza la acción.

Pregunta - ¿Cuál es la pirámide que está sobre el cubo que estaba en la caja ?

Para responder a esta pregunta, el robot consulta su "memoria", la respuesta producida es "la pirámide verde".

Pregunta - ¿ Hay algún bloque que soporte una más grande ?

Aquí se presenta otro tipo de referencia, la expresión "una más grande", quiere decir "una pirámide más grande que la pirámide verde".

Pregunta- ¿ Has movido algún bloque antes de haber cogido la pirámide azul ?

Le preguntamos por una acción, con una referencia al momento de su realización, primero tendrá que establecer el valor del tiempo T en el momento de coger la pirámide azul, después verificar si ha llevado a cabo algún desplazamiento de bloque antes de T.

La respuesta es : "El cubo rojo".

Pregunta - ¿ Cuándo ?

En este caso responde  $T=1$ , porque es el primer movimiento que hizo.

Pregunta - ¿ Por qué lo hiciste ?

Para responder debe determinar a que se refiere "lo", en este caso es el desplazamiento del cubo rojo, la contestación es : porque me lo ordenaron.

Este diálogo muestra la capacidad de comprensión de nuestro sistema, podemos afirmar que nos encontramos con un interlocutor que, en su medio, responde correctamente a las preguntas formuladas.

Aunque el vocabulario es restringido, hay una variedad en la construcción de las frases, con inclusión de pronombres; preguntas que implican un razonamiento, referencias al tiempo pasado, ect.

## REPRESENTACION Y TRATAMIENTO AUTOMATICO DE LA INFORMACION

Para que el robot pueda responder a las preguntas que se le formulan necesita no solo entender la lengua en la que le hablan, sino tambien tener información del medio que le rodea. Cuando se le ordena desplazar un objeto tiene que saber cómo hacerlo. Responder a cuestiones sobre el pasado implica memoria. Es decir, el proceso de comprensión lleva consigo una elaboración de la información recibida en la que intervienen conocimientos previos y métodos deductivos.

A la hora de escoger un lenguaje para representar la información y manejarla ~~adecuadamente~~; destacamos tres aspectos fundamentales a tener en cuenta para tomar una decisión:

- Primero, la posibilidad de representar información simple y compleja.
- Segundo, un método general de tratamiento, independiente del campo de aplicación.
- Tercero, que a la vez existan mecanismos que permitan indicar las características que en cada caso ayuden a mejorar los procesos de deducción.

PLANNER es un lenguaje general creado por HEWITT (1969), para la demostración automática de teoremas en el manejo de modelos en los que interviene un robot. En nuestro sistema hemos construido un lenguaje inspirado en PLANNER, que tiene los elementos necesarios para el modelo implementado, sin abarcar todas las posibilidades que el PLANNER ofrece, aunque es posible desarrollarlas.

La información que queremos representar consta de :

Objetos, sus propiedades, relaciones que mantienen entre ellos y con el robot. Podemos citar :

Clases de objetos : pirámide, caja, bloque, mano del robot.

Propiedades : manejable, libre, forma, color, tamaño...

Relaciones : posiciones respectivas, comparaciones..

Los objetos concretos los distinguiremos por llevar delante de su nombre el símbolo "\$", las propiedades y relaciones se representan precedidas de "\$".

La información simple se representa por medio de expresiones, por ejemplo, para definir que un objeto pertenece a una clase, escribiremos

( \$ ES :OBJETO 1 \$ BLOQUE )

Si queremos afirmar que es manejable :

( \$ MANEJABLE :OBJETO 1 )

Si se encuentra encima de otro objeto :

( \$ SOPORTA :OBJETO 2 :OBJETO 1 )

Mediante estas expresiones se describe la escena inicial en la que se encuentra el robot, para el caso del ejemplo dado en el apartado anterior tenemos el siguiente conjunto :

<u>OBJETOS</u>	<u>CODIGO</u>
Mesa	300
Caja	301
Piramide verde	302
Pirámide azul	303
Bloque rojo	304
Bloque azul	305
Mano del robot	306

( \$ ES :300 \$MESA )  
 ( \$ ES :301 \$ CAJA )  
 ( \$ ES :305 \$ BLOQUE )  
 ( \$ ES :304 \$ BLOQUE )  
 ( \$ ES :302 \$ PIRAMIDE )  
 ( \$ ES :303 \$ PIRAMIDE )  
 ( \$ ES :306 \$ MANO R )  
 ( \$ ES \$ ROJO \$ COLOR )  
 ( \$ ES \$ VERDE \$ COLOR )  
 ( \$ ES \$ AZUL \$ COLOR )  
 ( \$ ES \$ PUNTIAGUDO \$ FORMA )  
 ( \$ ES \$ RECTANGULAR \$ FORMA )  
 ( \$ COLOR :302 \$ VERDE )  
 ( \$ COLOR :304 \$ ROJO )  
 ( \$ COLOR :303 \$ AZUL )  
 ( \$ COLOR :305 \$ AZUL )  
 ( \$ FORMA :302 \$ PUNTIAGUDA )  
 ( \$ FORMA :303 \$ PUNTIAGUDA )  
 ( \$ FORMA :300 \$ RECTANGULAR )  
 ( \$ FORMA :301 \$ RECTANGULAR )



FACULTAD DE INFORMÁTICA  
BIBLIOTECA

( \$FORMA :304 \$RECTANGULAR )  
 ( \$FORMA :305 \$RECTANGULAR )  
 ( \$MANEJABLE :302 )  
 ( \$MANEJABLE :303 )  
 ( \$MANEJABLE :304 )  
 ( \$MANEJABLE :305 )  
 ( \$TAMAÑO :300 40 10 50 )  
 ( \$TAMAÑO :301 12 15 22 )  
 ( \$TAMAÑO :302 10 15 10 )  
 ( \$TAMAÑO :303 10 25 10 )  
 ( \$TAMAÑO :304 11 11 11 )  
 ( \$TAMAÑO :305 12 20 13 )  
 ( \$POSICION :300 60 20 30 )  
 ( \$POSICION :301 20 10 60 )  
 ( \$POSICION :302 50 10 40 )  
 ( \$POSICION :303 20 10 72 )  
 ( \$POSICION :304 20 10 60 )  
 ( \$POSICION :305 30 10 40 )  
 ( \$POSICION :306 20 10 30 )  
 ( \$LIBRE :304 )  
 ( \$LIBRE :305 )  
 ( \$LIBRE :302 )  
 ( \$LIBRE :303 )  
 ( \$SOPORTA :301 :304 )  
 ( \$SOPORTA :301 :303 )  
 ( \$SOPORTA :300 :301 )

La información más compleja se representa en forma de planes a seguir paso a paso, por ejemplo supongamos que queremos COGER un objeto que se encuentra sobre una mesa, las etapas a realizar serán las siguientes :

- 1- Probar que el objeto se puede coger ( para ello mirar si está definido como MANEJABLE ).
- 2- Probar que no hay otro objeto encima de él.  
 en caso de que lo haya, aplicar primero el plan "liberar el objeto X".
- 3- Verificar que no tenemos nada en las manos, si sujetamos algo, liberarnos de ello.
- 4- Desplazar la mano hasta donde se encuentre el objeto.
- 5- Cogerlo.

En PLANNER, esto se traduce en el programa siguiente :

```

THCON ( X Y ) ( $COGER Y X )

( THPROG ( X Y )
  ( THFIN ( $ ES X $OBJETO )
    THFIN ( $ MANEJABLE X )
    THFIN ( $ LIBRE X )
    THUTILIZAR QUITAR
  )
  THY
  ( THFIN ( $ ES Y $MANO )
    THFIN ( $ LIBRE Y )
    THUTILIZAR LIBERAR )
  THFIN ( $DESPLAZAR Y )
    THUTILIZAR DESPLAZAR
  THANADIR ( $COGER Y X ) ) )

```

En la primera línea aparece la expresión ( \$COGER Y X ) precedida de THCON ( X Y ), que quiere decir : este es el plan a seguir para comprobar a partir del conjunto inicial de expresiones que describen la escena , si la acción se puede llevar a cabo o no.

Estos programas se construyen siguiendo las reglas sintácticas :

```

PROGRAMA ----- PP1 / PP2
PP1 ----- THPROG ( VARIABLE ) ( CUERPO )
PP2 ----- THBUSCAR ( PARAMETRO ) ( VARIABLE ) ( CUERPO )
PARAMETRO----- entero cons signo
CUERPO ----- EXPRESION COMPUESTA / EXPRESION COMPUESTA ( PROGRAMA )
                EXPRESION COMPUESTA ( PROGRAM) ( CUERPO ) / EXP.COMP ( CUERPO )
EXPRESION COMPUESTA -- OPERADOR ( ( EXP SEC ) ( SUCESION EXP SEC ) ) /
                THNO ( EXP SEC ) / EXP PRIMARIA

```



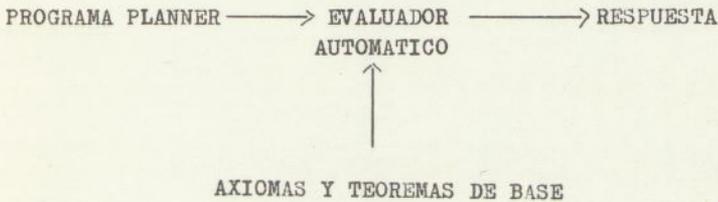
Si entre la base de axiomas tenemos :

\$ES :OBJETO 1 \$BLOQUE

\$COLOR :OBJETO 1 \$AZUL

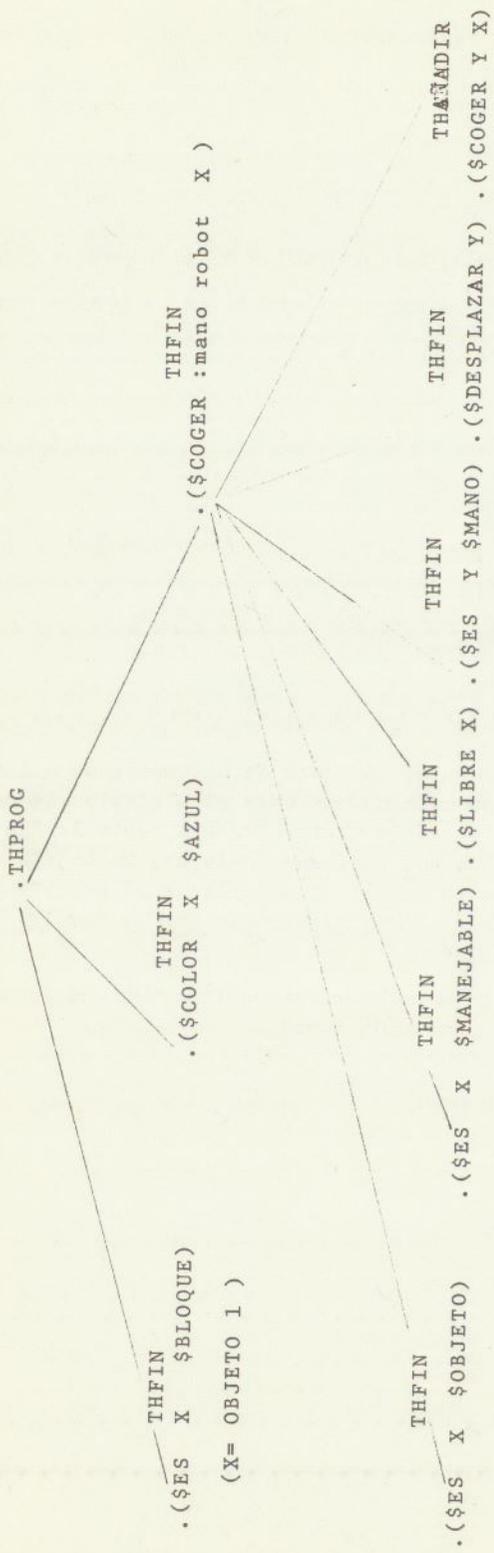
Y en la base de teoremas THCON ( X Y ) ( \$COGER Y X )

El demostrador, que es en realidad un EVALUADOR de programas PLANNER desarrolla la arborescencia de la página siguiente en un proceso que podemos representar por el diagrama :



En la primera hoja compara con el conjunto de axiomas y da a X el valor :OBJETO1 ( la comparación se hace mediante el algoritmo de unificación ), con esta elección prosigue verificando el resto de las ramas. En este ejemplo simplificado, en el que solo hay un objeto que responde a las características no es necesario aplicar al mecanismo de "vuelta atrás" para reconsiderar las elecciones que se revelan incorrectas posteriormente.

Este mecanismo lleva una serie de heurísticos incorporados para hacer lo más eficaz posible la decisión tomada.



En el nodo \$COGER se inserta la arborescencia correspondiente al teorema COGER.

La frase "Pon la pirámide azul sobre el bloque más grande" después de ser analizada es el siguiente programa :

```

THPROG ( X Y )
  ( THFIN ( $ ES X $ PIRAMIDE )
    THFIN ( $ COLOR X $ AZUL )
    THBUSCAR ( 1 ) ( Y )
      ( THFIN ( $ ES Y $ BLOQUE )
        THNO
          ( THEUSCAR ( 1 ) ( Z )
            ( THFIN ( $ ES Z $ BLOQUE )
              THFIN ( $ MAS $ TAMAÑO Z Y )
                THUTILIZAR ( MEDIDA ) ) ) ) ) )
    THFIN ( $ PONER SOBRE X Y )
    THUTILIZAR ( PONER ) )

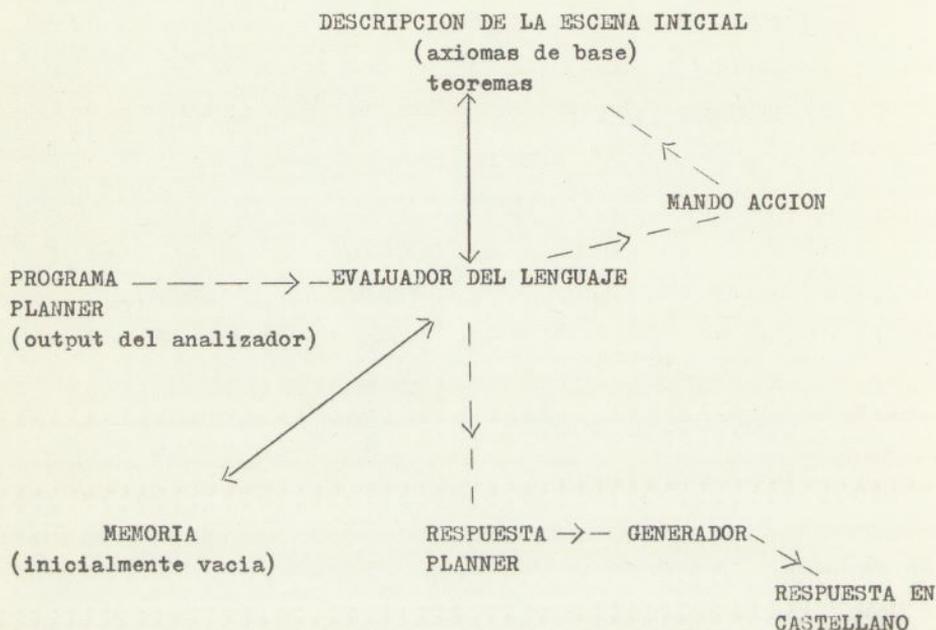
```

La evaluación de este programa se hace superando las etapas :

- Determinar un objeto (dar un valor a X) tal que pertenece a la clase pirámide y verifique que sea azul.
- Determinar un objeto (dar un valor a Y) tal que pertenezca a la clase bloque y verifique que no haya otro objeto Z de la misma clase que sea más grande que él.
- Llevar a cabo el plan de base PONER para realizar la acción.

Si se acaba con éxito se genera una respuesta positiva, dando las nuevas coordenadas del objeto desplazado, sino se ha realizado se contesta explicando el por qué.

Esta etapa del sistema se resume mediante el siguiente organigrama :



Una vez establecida la representación interna y definidos los métodos de tratamiento de la información, el problema que hay que resolver es dada una frase en castellano, obtener su representación en PLANNER.

Este proceso lo realiza un analizador sintáctico y semántico.

FRASE EN CASTELLANO —> ANALIZADOR —> PROGRAMA PLANNER