

## EL PROYECTO MMI<sup>2</sup>: LA CONSTRUCCION DE UN INTERFAZ MULTIMODAL PARA INTERACCIONES HOMBRE/MAQUINA

Ricard Pérez  
Departamento de I+D / NLU de ISS,sa

**Resumen.**- El artículo presenta algunas de las ideas que se implementan en un interfaz para diálogo multimodal hombre / máquina con sistemas expertos que se desarrolla en el proyecto MMI<sup>2</sup> (ESPRIT nº 2474). Presenta sus objetivos y un perfil de la arquitectura global y de las funcionalidades de los módulos que la constituyen. El artículo hace especial hincapié en la caracterización del módulo de lengua española (SPM) del interfaz<sup>1</sup>.

### Introducción

Este artículo presenta aspectos del interfaz que se desarrolla en el proyecto MMI<sup>2</sup> dentro del marco del Programa comunitario ESPRIT II (proyecto nº 2474: MMI<sup>2</sup>: "Multi-modal Interface for Man-Machine Interaction"). El proyecto reúne una treintena de investigadores especializados<sup>2</sup> en diversos campos de la IA conectados a diversas instituciones<sup>3</sup>, dirigidos por J-L Binot, director científico de la sociedad belga BJM,sa.

### 1. Objetivos del proyecto MMI<sup>2</sup>

El objetivo del proyecto MMI<sup>2</sup> es construir un interfaz adaptable a distintas tipologías de usuarios, que ofrezca, dentro de un marco unificador del diálogo, diversos modos integrados de expresión de entrada y salida, para servir a una comunicación robusta de tareas cooperativas compartidas entre el usuario y el sistema, con sistemas de bases de conocimiento experto. El proyecto se desarrolla en un entorno de potentes estaciones de trabajo Sun/3 y Sun/4, y el núcleo del interfaz se implementa en BIMproLog.

Veamos más atentamente algunos puntos que hemos resaltado en la frase anterior:

1. Interfaz.- Definir el objetivo global como la construcción de un interfaz equivale a concebir el prototipo MMI<sup>2</sup> resultante en oposición a una concepción del mismo como mero 'front-end' de una aplicación concreta. El proyecto debe resultar por tanto en una delimitación de los cambios a realizar en el interfaz cuando deba conectarse a distintos SBC ('Sistemas de Bases de Conocimiento'), siendo constante del proyecto la atención a principios metodológicos en los distintos diseños para facilitar una definición clara de las líneas de adaptación a nuevas aplicaciones (modularidad, conocimiento dependiente-independiente del modelo del mundo, del dominio de aplicación, de las tareas cooperativas genéricas o específicas, etc.).

<sup>1</sup> El módulo SPM junto con el módulo de la semántica multimodal (SE) y el módulo de planificación de la comunicación (CPM) se desarrollan en el departamento de I+D /NLU de ISS, sa.

<sup>2</sup> Entre los investigadores del equipo de ISS,sa. participantes en el proyecto MMI<sup>2</sup>, bajo la dirección del autor están J. Barreras M.Gatius, A. Manzanera, M.A. Pérez, J.C. Ruiz y D. Trozig.

<sup>3</sup> En consorcio técnico del proyecto MMI<sup>2</sup> está formado por investigadores que colaboran en BIM,sa (Everberg) dirigidos por J.-L. Binot; en el departamento de I+D/NLU de ISS,sa (Barcelona) por R. Pérez; en el departamento de Psicología del INRIA (Domaine du Voluceau-Rocquencourt), bajo la dirección de P. Falçon; en EMSE (Saint-Etienne), por B. Peroche; en el CRISS de la Universidad de Grenoble (Grenoble), por J. Rouault; en RAL (Oxford) por R. Ringland y M. Wilson, y en el departamento de Psicología y Ergonomía de la Universidad de Leeds (Leeds) por M. Howes y N. Sheehy, vid [J.-L. Binot et al 1991]

2. Tipologías de usuario.- Definir tipologías de usuarios permite al interfaz MMI<sup>2</sup> adaptar el interfaz al comienzo de la sesión de trabajo y, también, de forma dinámica, en el transcurso de una sesión y en el transcurso de varias sesiones, de acuerdo con determinados ejes que configuran las tipologías. Este conocimiento está definido en el Experto de Modelación del Usuario (UM). La inferencia principal se hace sobre clases de usuarios, sea su nacionalidad (para decidir el lenguaje natural para la comunicación del interfaz, las unidades de cuenta preferentes,...), sistemas de preferencia personalizados (modo preferencial de generación de respuestas, etc.), experiencia en el manejo del interfaz (principiante, experto), interés profesional preferencial (cliente, ejecutivo, técnico, experto en la aplicación), registro de la dinámica de comunicaciones de asesoramiento al usuario dentro de la sesión, etc.

3. Marco unificador de diálogo.- Usar un marco unificado de diálogo conlleva integrar en un registro unitario todos los acontecimientos que acaecen en el interfaz. Este objetivo ha buscado consolidarse mediante un registro del diálogo multimodal dentro de un marco amplio del discurso y su contexto, un lenguaje formal de representación común para todos los modos de expresión del interfaz - el formalismo CMR -, un conjunto de etiquetas de un Experto Semántico (SE) que configura el mundo del interfaz MMI<sup>2</sup> - etiquetas de representación lingüística y descripción objetual [Hobbs 1985], [Hirst 1987] - y una gestión unificada del diálogo.

4. Modos integrados.- Los modos son medio de expresión y se distinguen de los medios. Por ejemplo, un 'medio' gráfico de comunicación puede ser soportado por diversos modos gráficos, tal como ocurre en el proyecto MMI<sup>2</sup>. Los modos característicos para la entrada de comunicaciones al interfaz MMI<sup>2</sup> adaptables a la aplicación son los gestos, los comandos, diversos modos gráficos, los lenguajes naturales (inglés, francés y castellano) con o sin deíxis gráfica; y para la salida de información, diversos modos y facilidades gráficas de respuesta, paráfrasis, si cabe, del lenguaje natural en comandos, los lenguajes naturales (inglés, francés y castellano) con o sin deíxis gráfica.

El concepto de unificación se complementa además con la mezcla de modos en una misma expresión, por ejemplo, deíxis gráfica dentro de los LNs y comandos, la referencia en el LN a objetos del propio interfaz y a propiedades de atención visual, la paráfrasis del LN por comandos, el empleo de sintagmas nominales en el interior de comandos, etc.

5. Comunicación robusta.- Dada la cantidad y variedad de niveles en los que potencialmente pueden aparecer fenómenos de interferencia e interrupción de transmisión informativa en el entorno de una comunicación dialogada, la robustez de la misma constituye un proyecto ambicioso en sí mismo. La robustez, como conjunto de medios de recuperación del diálogo ante incidencias tendentes a la distorsión e interrupción de la comunicación hombre-máquina, supone enfrentarse a situaciones muy diversas cuya caracterización y cometido de recuperación depende en gran medida del diseño concreto de herramientas y sistemas, así como de la sagacidad e inteligencia de los mismos. Este preámbulo nos sirve para decir que el proyecto MMI<sup>2</sup> no ha planteado dirigir sus esfuerzos a la investigación de la robustez en general sino a la que se deriva necesariamente del diálogo multimodal, sirva como ejemplo, la ambigüedad, la incertidumbre o el fallo ante presuposiciones referenciales, o la incapacidad de cometido inmediato del sistema de aplicación. El método general focalizado en ambos casos es la creación, a iniciativa del sistema, de planes de comunicación subdialogados usando tipologías argumentales, para la reparación del estado del diálogo. Temas como la corrección de errores tipográficos, la agramaticalidad, la incoherencia, etc. son dejados a cargo de los modos de entrada sin acceso a la comunicación con el interfaz nuclear.

6. Tareas y diálogo cooperativo.- Siguiendo el hilo del último párrafo, cabe mencionar que el proyecto focaliza los diálogos cooperativos en tareas compartidas de diseño como actitud general de los usuarios del interfaz. La iniciativa de usuario se plantea como eje del diálogo

principal, mientras que el sistema toma la iniciativa preferentemente en los subdiálogos. Se propone la tutoría en el aprendizaje del uso del propio interfaz, así como se hace el supuesto de una actitud de adaptación del usuario a la carga conceptual y a la complejidad expresiva exhibida por el propio sistema, esperándose así mismo, en gran medida, el cumplimiento de las máximas cooperativas de Grice [1975].

7. Sistemas basados en conocimiento. - El segundo gran objetivo metodológico del proyecto es determinar las características de los sistemas expertos conectables al interfaz MMI<sup>2</sup>. Los sistemas de conocimiento para el interfaz MMI<sup>2</sup> están basados en Prolog y, en particular, el prototipo del propio proyecto MMI<sup>2</sup> funciona con un sistema de bases de conocimiento (SBC) especializadas en el análisis y diseño de redes de ordenadores, que ha sido desarrollado dentro del mismo proyecto. Este tema no es abordado, sin embargo, en este artículo.

## 2. El formalismo y el léxico conceptual común de representación

Para hacer posible la arquitectura que implemente esta integración de funcionalidades, se ha creado un formalismo de representación semántica común a todos los módulos, que se usa como vehículo de comunicación interna dentro del interfaz, y empleado también como base de razonamiento e interpretación semántica y pragmática. Este formalismo se llama CMR (Common Meaning Representation).

Consiste en una lógica de predicados de primer orden potenciada con diversas anotaciones sobre las variables y otros niveles de escritura.

Las fórmulas de las expresiones de CMR contienen predicados reificados, esto es, una lógica en la que los nominales, adjetivos, sintagmas verbales y adverbios pueden representarse igual como los nombres, es decir, de los cuales podrá afirmarse o negarse su existencia. Reificación proviene del latín 'res', cosa, objeto, y, efectivamente, la reificación consiste en hacer cosas, objetos, de elementos que en apariencia no lo son, como puede ser un acontecimiento (por ejemplo, una 'explosión' es comparable a una 'montaña', podemos hablar de sus perímetros, etc.).

Veamos un poco más de cerca el concepto de reificación en sus determinaciones para la lógica. Una fórmula lógica para representar una frase de lenguaje natural como 'Eva come la manzana' podría ser la siguiente:

(1) comer (Eva, manzana),

la forma lógica de (1) se dice que es una fórmula singular, y como podemos ver contiene un predicado comer /2, el cual podría ser cierto en un mundo-modelo en el que sólo existiesen dos objetos en su dominio, es decir, 'Eva' y 'la manzana'. La dificultad de esta representación lógica es que si añadimos un adverbio a nuestra frase, por ejemplo, 'Eva come gustosamente la manzana', obtendríamos (1')

(1') comer\_gustosamente (Eva, la manzana)

con un nuevo predicado comer\_gustosamente/2 completamente imposible de equiparar al anterior comer /2. Como los adverbios y las frases adverbiales pueden multiplicarse y concatenarse sin fin, el análisis lógico debería abdicar ya que cualquier modificación adverbial daría lugar a un predicado distinto e irreducible. La consecuencia lógica es también absurda, pues se plantea una representación del lenguaje con un léxico básico sin límite.

De otro lado, también (2) representa la frase original:

(2) (Ex) (comer (x) & agente(x, Eva) & paciente (x, manzana))

(2) es una fórmula existencialmente cuantificada, ahora la fórmula contiene tres predicados, *comer/1*, *agente/2* y *paciente/2*; esta diferencia que puede parecer insignificante representa un cambio substancial en lógica ya que la frase sólo puede ser cierta en un mundo-modelo donde por lo menos existan cinco objetos: Eva, la manzana, la comida de Eva, la comida de la manzana y la comida de la manzana por Eva. La extensión de *comer/1* tiene que ser por lo menos 'la comida de Eva' y 'la comida de la manzana'. Vemos así como *comer/2* se ha transformado en *comer/1*, más *agente/2* y *paciente/2*; ontológicamente se ha extendido el dominio de dos a cinco objetos. En lógica se dice que estamos cuantificando sobre acontecimientos o lo que es lo mismo, reificando los acontecimientos. Ahora los adverbios pueden acomodarse:

(2') (Ex) (comer (x) & gustosamente (x) & agente (x, Eva) & paciente (x, manzana))

Asimismo, el formalismo hace recurso de estructuras lógicas con un número canónico de argumentos para representar predicados verbales, que de otro modo tendrían un número de argumentos variable dependiendo del número de complementos regidos por el predicado.

Las etiquetas conceptuales del Experto Semántico del interfaz (SE), que equivale al léxico conceptual del lenguaje CMR, responden justamente a las reificaciones de esos predicados<sup>4</sup>.

Las variables y las constantes de las fórmulas pertenecen a los tipos de las etiquetas del SE. Es decir, la CMR es un formalismo que sigue las representaciones lógicas normales del tratamiento del lenguaje natural en lo que se llama lógica de tipos.

Una frase como 'Todos los hombres son mortales' pasaría de tener una representación como

(3)  $(\forall x) (\text{hombre}(x) \rightarrow \text{mortal}(x))$ ,

a una representación tipificada como

(3')  $(\forall x:\text{hombre}) \text{mortal}(x)$

Desde nuestro punto de vista podemos considerar esta 'tipificación' lógica como una conveniencia notacional que permite a la sintaxis de la CMR parecerse más a la sintaxis del lenguaje natural.

Otro nivel de la CMR son las anotaciones de las variables, entre las que se incluyen las de definido-indefinido, singular-plural, pronombre, anáfora,... Como puede apreciarse dichas anotaciones tienen un claro sentido para la resolución de problemas referenciales, pero también las hay que no se refieren tan directamente a estos fenómenos, como el tiempo verbal, si hay, la tipología de la emisión (interrogativa polar, imperativa, aserción,...).

Otros niveles de escritura se refieren al tiempo interno de la emisión, el identificador del módulo que ha creado la CMR, el estatus del diálogo, malas concepciones, etc. En realidad, el proyecto en curso ha explotado hasta ahora sólo algunas de las previstas.

---

<sup>4</sup> Hay un premio adicional a la reificación de acontecimientos y adverbios dentro de la estructura clara y elegante de la lógica de predicados clásica, ya que ésta admite funciones de verdad, cuantificación y predicados de uno o más lugares con variables vinculadas. La identidad puede introducirse como uno más de los predicados de dos lugares ('aridad/2'), y los términos constantes singulares y signos de función son fácilmente parafraseables, en el contexto, para que entren en el esquema. Esta lógica es susceptible de procedimientos de prueba conocidos cuya completitud ha estado extensamente demostrada. Gracias a Tarski, la estructura contiene una definición recursiva de satisfacción y verdad.

### 3. Descripción general del interfaz MMI<sup>2</sup>

La arquitectura de MMI<sup>2</sup> es modular. El sistema está formado por módulos expertos especializados en grupos coherentes de tareas. Los módulos se interrelacionan por medio de predicados públicos, escritos en Prolog, que responden a las significaciones informativas que cada módulo pone a disposición de los demás módulos. Por tanto, la semántica transmitida por los predicados públicos orienta a los diseñadores de otros módulos y debe resultar en la transmisión conceptual intermodular necesaria de unos con otros.

En MMI<sup>2</sup> se atiende, ante todo, a aquellos aspectos que inciden en la integración de la comunicación multimodal y la comunicación con las bases del conocimiento del SBC de aplicación.

No obstante, otros aspectos están siendo también atendidos a diferentes niveles de profundidad. Aspectos tales como la robustez, los subdiálogos de reparación, las desambiguaciones conceptuales, o de acepciones léxicas, etc. aparecen entre los 'scripts' que se usan como 'lecho de tests' de los avances en la comunicación soportada por el interfaz.

#### 2.1. Modos expertos

Los modos del interfaz representan las formas genéricas de manipulación disponibles para usuario y sistema. Debemos distinguir los modos expertos, de las formas que pueden acoger las interacciones concretas. Las interacciones pueden acoger formas de expresión mixtas, es decir, una misma interacción puede conllevar la manipulación de dos modos distintos (no hay previsión de manipulaciones mixtas de más de dos modos). Un ejemplo de forma mixta de interacción es, por ejemplo, el empleo de lenguaje natural y de la deíxis gráfica en una misma intervención.

Aquí, sin embargo exponemos las formas de manipulación, es decir, los módulos que gestionan un tipo simple de manipulaciones de entrada/salida del interfaz.

**Gesto.-** El modo experto de manipulaciones gestuales. Obviamente es un modo únicamente de entrada. El modo ha sido construido de manera que fácilmente puedan construirse nuevos gestos que se conecten a nuevas funcionalidades. En realidad, el gesto actúa como un pre-modo, en el sentido que no es él mismo quien se encarga de proyectar el contenido proposicional resultante de los gestos esgrimidos por el usuario sino que lo es el modo al cual afecta la manipulación gestual. De forma generalizada, los gestos en los prototipos MMI<sup>2</sup> actúan sobre los modos gráficos.

**Gráficos.-** Los modos expertos de manipulación gráfica. De momento se ha desarrollado un modo gráfico para trasladar planos de edificios, otro, de estrecha relación con el anterior, para plasmar gráficamente diseños de redes de área local para esos edificios, y, también, una serie de modos gráficos consistentes en tablas, gráficos de barras, pasteles,... Los dos primeros modos permiten el dibujo a mano alzada con ratón, por parte del usuario, así como una corrección normalizada o recomposición automática del dibujo por parte del gestor gráfico. La comunicación gráfica del diseño se hace a través de los llamados 'planos planos' ('planar maps'), es decir, una sucesión de cortes horizontales para la representación de las plantas de los edificios. Estos modos aceptan la superposición transparente de unos con otros así como el 'zoomming', en el cual, el espacio visualizado por pantalla en cada momento de la sesión se gestiona como un 'foco' (visual) para el propósito de la comunicación hombre/máquina. Estos modos admiten deíxis y referencia plural (de conjuntos de objetos pero no de clases de objetos). Los otros modos gráficos se utilizan preferentemente para la generación de respuestas por parte del sistema, aunque, sin embargo, soportan también la deíxis por parte del usuario.

Un gestor gráfico proyecta las interacciones gráficas en el formalismo común CMR y viceversa.

Comandos.- Evidentemente, se trata de un modo de comunicación imperativa con sintaxis convencional muy constreñida.

El propósito del lenguaje de comandos es suministrar al usuario una gran eficiencia imperativa a costa de tener que trabajar con una sintaxis restrictiva. El interfaz intenta explotar aspectos de la arquitectura multimodal para facilitar una transición desde el lenguaje natural imperativo al lenguaje de comandos. El lenguaje de comandos dispone de una ventana exclusiva para comandos sobre la presentación visual de la pantalla. Esta solución ha sido adoptada al no existir, de momento, un test claro para distinguir entre lenguaje natural razonablemente bien formado y lenguaje de comandos mal formado.

El lenguaje de comandos tiene una parte sintáctica y otra semántica, la primera reside en el Modo comandos, la segunda en el Experto Semántico. La parte sintáctica comprenden dos conjuntos de datos, un conjunto de operadores/acciones y el segundo un conjunto de objetos del mundo del interfaz MMI<sup>2</sup>. La sintaxis de superficie de los comandos exige el uso de un par 'acción-objeto'. Los pares 'acción-acción' y 'objeto-objeto' no son admitidos. Para determinar si un comando se conforma a la semántica de la aplicación, los comandos se codifican en CMR y se evalúan por medio de las constricciones residentes en el modelo del mundo de MMI<sup>2</sup> del Experto Semántico.

Según su funcionalidad, el lenguaje de comandos comprende: 1- metacomandos (macros,...), 2- operaciones de ficheros (salva, carga,...), 3- comandos específicos del usuario (alias, lenguaje, alertas,...), 4- comandos sobre el diálogo (historia, vuelve a hacer, redibuja, deshaz, salva estado, restaura estado, corta,...) 5- Comandos de búsqueda de información (ayuda, cómo, porqué, justifica,...), 6- acciones gráficas (limpia, mueve, añade, reemplaza,...) , 7- comandos de diseño (chequea, evalúa, lista, computa, conecta, desconecta,...)

Lenguajes Naturales.- Los modos expertos de traducción del lenguaje natural en una representación en el formalismo común CMR, y viceversa, desde representaciones CMR a las formas superficiales de cada uno de los lenguajes naturales presentes en el interfaz: Español, Francés e Inglés. La comunicación mediante los modos de L.N. no pueden alternarse de intervención a intervención, es decir, el diálogo transcurre con la participación de uno solo de ellos; por tanto, no debe pensarse en un diálogo multi-lingüístico, el cual requeriría conceptualizarse como una traducción automática con un interlingua.

## 2.2. Módulos expertos

Experto semántico (SE).- Es un fondo conceptual común de etiquetas simples para las representaciones CMR de todos los modos. Tiene conocimiento del modelo del mundo de MMI<sup>2</sup>, esto es, un cierto conocimiento intensional de los prototipos de los objetos de su mundo conectado a una aplicación concreta. Está constituido básicamente por una jerarquía que refleja cuatro metaniveles de clasificación: tipos, roles, roles evaluativos y roles de roles evaluativos. Estos metaniveles permiten trasladar al interior del interfaz ciertas ambigüedades de significado de una misma etiqueta conceptual. Su conocimiento se usa también para ciertos problemas de referencialidad. Contiene el conocimiento semántico de los comandos del interfaz. También el Experto en el Dominio (DE) forja los conceptos complejos (por ejemplo, 'salas\_adyacentes', 'coste\_del\_server', etc.) necesarios para la aplicación, a partir de sus conceptos simples. Contiene asimismo la conceptualización de los diagnósticos generales sobre malos funcionamientos, disfunciones, etc. con los que el Experto en Planes de Comunicación construye diálogos argumentativos de reparación.

Experto en el contexto (CE).- Conecta con el Resolvedor de anáforas intersentenciales que es uno de sus constituyentes, gestionando la propuesta de candidatos de resolución. Tiene a su cargo todo lo referente a identificar la estructura del diálogo y sus diversos componentes:

registrar esa estructura, extraer información relevante, gestionar los 'focos' y 'tópico' del discurso.

Control del impacto formal del diálogo (DC).- Gestiona la estructura del diálogo, activa los expertos necesarios para mantener las interacciones del mismo. Su interés se centra en el impacto formal de las intervenciones. La idea es que la validez formal de la entrada no es suficiente para garantizar el cometido pragmático, ya que se necesita también una evaluación del impacto informal de la comunicación. Así pues, el trabajo de control se divide en dos y el DC toma a su cargo sólo la evaluación del impacto formal. El control formal del diálogo se realiza a través de una pila LIFO que acumula las propuestas de subdiálogos a comunicar de cada uno de los módulos que evalúan el impacto informal de las intervenciones. En el momento en que el subdiálogo debe ser generado, el DC recurre al planificador de comunicaciones. Las contrapartidas de respuesta por parte del usuario, cuando las hay, se dirigirán en especial a la evaluación del módulo proponente (expectativas,...). El impacto informal de la comunicación es evaluado por los módulos CE, UM, DE e IE.

Modelo del usuario (UM).- Mantiene y explota las tipologías de usuarios. Selecciona el modo de lenguaje natural, unidades de cuenta, etc. al principio de la sesión. Asesora en la elección de modos de respuesta según las preferencias del usuario y su estatus de conocimientos del propio interfaz y del dominio de aplicación. Asesora sobre el tipo de planes de comunicaciones a comunicar al usuario en caso de subdiálogos de reparación. Registra las comunicaciones habidas durante la sesión corriente para evitar redundancias inmediatas en planes sucesivos de comunicación, etc.

Modelo del dominio y las tareas (DE).- Tiene conocimiento y gestiona los planes de tareas cooperativas potenciales del SBC para describirle los problemas que provienen de las comunicaciones del usuario, conoce los conceptos complejos con los que necesita trabajar el SBC y los forma a partir de la expresión CMR, comunica al CPM los diagnóstico del SBC.

Modelo del propio interfaz (IE).- Una de sus funcionalidades es la de mantener la gestión de comunicaciones del interfaz con los elementos del entorno de la estación de trabajo, Sunviews y sistema operativo, es decir, gestionar los acontecimientos que ocurren en las ventanas suministradas por el interfaz. Como experto en el propio interfaz MMI2, sólo se ha desarrollado de forma incipiente. Teóricamente debe tomar a su cargo el aspecto del conocimiento tutorial del propio interfaz.

Módulo de planes de comunicación del sistema (CPM).- Planifica las respuestas del sistema teniendo en cuenta el nivel intencional de la función argumentativa de cada contenido proposicional generado por el sistema. Esquema del Interfaz MMI2 La aplicación conectada al prototipo actual consiste en un KBS de análisis y diseño de redes de área local.

El esquema de la página siguiente presenta una visión de la estructura global del sistema

#### **4. Aproximación dinámica al funcionamiento del interfaz**

Este apartado intenta una animación muy sintética y sencilla del interfaz MMI<sup>2</sup>. Para situar la redacción nos hemos valido de la numeración y siglas que figuran en el esquema del interfaz.

En el centro del esquema vemos en líneas discontinuas la representación de una comunicación, en el formalismo <Common-Meaning-Representation> (CMR) que denominamos un 'CMR\_act\_analysis'. Esta representación está compuesta de la sintaxis del formalismo CMR más un contenido proposicional que se expresa a través de tipos legalizados por el SE (- 3 -). En principio puede provenir de cualquiera de los modos expertos de entrada (- 2 -) proveniente de un acontecimiento de interacción del usuario sobre el interfaz (- 1 -), el cual ha sido gestionado por el IE (en el esquema no se ha representado esta funcionalidad del IE).

De contener pronombres o anáforas intersentenciales, el CE y su resolvidor de anáforas (- 4 -) decidirán el candidato mejor colocado de la referencia, poniendo en juego informaciones diversas como son: información sintáctica dependiente del modo L.N. (si la información proviene de un modo L.N.), información de focos y del tópico que gestiona el propio CE, información conceptual del SE, etc. La CMR es entonces registrada y los focos y tópico actualizados.

El DC en base a las anotaciones sobre los tipos de las expresiones o 'U\_types' (interrogativa, imperativa, asertiva, etc.) y al estado del diálogo y con el asesoramiento del CPM, asigna un impacto formal ilocucionario o fuerza comunicativa, y seguidamente da juego a los módulos que deben evaluar el impacto informal: UM, IE, DE, los cuales recogen las partes interesantes de la CMR para sus registros y, en el caso de decidir comunicaciones subdialogadas, las envían a una pila LIFO del DC, (- 7 - 8 - 9 -).

Si el DE después de recoger los conceptos complejos necesarios para el KBS y los datos de la tarea o problema está en disposición de transmitirlos al KBS, proyectará al lenguaje del KBS la estructura resultante acumulada por la comunicación. El KBS hace su cometido devolviendo su respuesta al DE (- 10 - 11 -). Si la respuesta es directa, el DE se encarga de la proyección a CMR, sino comunica con la pila LIFO del DC para depositar su demanda de subdiálogo, pasando los correspondientes diagnósticos e informaciones necesarios para la creación, en su momento, de la instancia del plan de subdiálogo (- 12 -).

De cualquier forma, el DC decide sobre el impacto formal de la acción de retorno y da orden al CPM de establecer un plan de comunicación con el usuario (- 13 -) que con el asesoramiento de informaciones del UM decidirá los modos expertos que intervendrán en la generación de la comunicación, etc.

El CPM crea su plan de argumentación consistente en CMRs (- 14 -), que serán registradas por el CE, y deposita las expectativas derivadas de su plan de interacción para el control del DC y los demás módulos de gestión del diálogo.

Las CMRs resultantes pasan a los generadores de los diversos modos de salida (- 15 -) que finalmente es comunicado al usuario (- 16 -).

## **5. El modo de expresión en lengua española: SPM**

### **5.1. Funcionalidades**

La funcionalidad del SPM en conjunto es la de analizar la entrada del usuario construyendo una representación funcional de acuerdo con la sintaxis expresada por el formalismo CMR y las etiquetas del Experto Semántico del interfaz MMI2.

El énfasis del diseño del SPM se ha puesto principalmente en cuatro ejes:

1. La flexibilidad en la administración y guía del modo a diferentes aplicaciones, dominios y fuentes lingüísticas, es decir, favorecer el eje portabilidad-adaptabilidad: interfaces adecuadas para los módulos, con recursos adecuados para dar de alta y de baja los conocimientos, cambiar datos, etc.;

2. El eficiente rendimiento en el run-time: programación en C del analizador morfológico, gestión del tamaño del léxico de run-time, reducción del backtracking en el parsing, administración de la ambigüedad de significados de los lemas, cooperación con módulos del interfaz nuclear: entre otros, la colaboración con el DE para conocer sobre nombres propios de individuos, colaboración con el SE para el vinculamiento de frases preposicionales, decidir

sobre ambigüedades estructurales, en fenómenos de coordinación, en la expresión de semánticas implícitas, sean genéricas del mundo de MMI<sup>2</sup> o del dominio de aplicación, en la especificación de las relaciones de significado entre entidades lingüísticas, etc.);

3. La independencia modular de los componentes y la fijación de las fronteras y estructuras de integración de los componentes, permitiendo ya sea la substitución de módulos, ya sea la funcionalidad independiente, ya la posibilidad de entrar en otras configuraciones de L.N. También es importante para conservar la independencia modular el control de los fenómenos lingüísticos dentro de los módulos del SPM y de las interrelaciones en su solución. Asimismo tomamos en consideración los diseños que permitiendo la intervención de aspectos dominio dependientes (por ejemplo, semántica del dominio) tuvieran, sin embargo, posibilidad de soslayarse si se usan en entornos menos restrictivos.

4. La robustez y la recuperación de errores en el reconocimiento del L.N.: básicamente frente a los fenómenos de palabras no reconocidas, entradas no parseables, o, limitaciones de la expresividad del formalismo de representación, no obstante, también estamos estudiando implementar técnicas de aprendizaje dinámico de conocimientos.

Los planes de extensión del SPM enfocan entre otros, la creación de su propio interfaz integrado dentro de un marco interactivo para que un administrador de lenguaje español pueda de forma fácil y flexible (ergonómica) guiar los recursos del SPM a los diferentes dominios y fuentes lingüísticas manteniendo la consistencia conjunta de las informaciones presentes en las herramientas.

Como el interfaz MMI<sup>2</sup> aspira a un reconocimiento y generación de un lenguaje natural que sea 'natural', el SPM transfiere al interfaz parte de las ambigüedades de significado de los lemas, los problemas de referencia intersentencial al interfaz nuclear, ya que dependen del contexto en el que las expresiones son emitidas. En su caso, el SPM se limita a pasar información sintáctica lenguaje dependiente para asesorar la resolución de las anáforas que tiene a su cargo el CE.

Los componentes del SPM actual son los siguientes (véase el esquema del SPM en la página siguiente para apreciar la situación y relaciones de cada módulo):

- Pretratamiento morfológico
- Analizador morfológico: MORFEO
- Filtro
- Léxico Virtual
- Léxico Actual
- Parser
- Herramienta de paso al formalismo CMR

### 5.3. El pretratamiento morfológico

El pretratamiento morfológico toma a su cargo básicamente una cierta regularización de puntuaciones de las cadenas y desdobra las contracciones al y del.

### 5.4. El analizador morfológico: MORFEO

El analizador intenta analizar los nombres propios como nombres comunes (clases clónicas, como por ejemplo 'estación Sparc'), si no lo consigue, consulta el tipo semántico de su denotación extensional al Experto en el Dominio (DE), de obtener una respuesta negativa por parte del DE, prevemos implementar técnicas inteligentes de robustez y en su defecto, mensajes al usuario, por parte del propio modo SPM, que notifiquen el desconocimiento del interfaz, evitando de esta forma la ruptura de la comunicación después de un mayor tiempo de procesamiento.

El conocimiento estático del analizador morfológico consiste en un diccionario de bases estructurado en un árbol de letras y una modelación de sufijos y desinencias. El carácter blanco es reconocido como un carácter más, lo cual favorece la potencia de la herramienta para el reconocimiento de lexías, así como su adaptabilidad a expresiones conceptuales estereotipadas de los dominios de aplicación. La estructura del diccionario de bases proporciona un sólido fundamento para dotar a la herramienta de un mecanismo potente de resolución de errores tipográficos<sup>5</sup>.

El perfil básico del proceso de análisis morfológico consiste en reconocer la base de mayor longitud posible para ver si con el resto pueden reconocerse los modelos válidos de sufijo y desinencia para esa base. Si es así, el análisis para esa palabra consolida las informaciones estáticas y dinámicas de la base y modelos aplicados. De nuevo vuelve a intentar una segmentación de menor longitud de base, a no ser que la base reconocida en el análisis anterior conllevara una marca de indivisibilidad, en cuyo caso, el analizador declina segmentaciones sucesivas de la palabra y captura una nueva forma superficial de la cadena de entrada.

### 5.5. El filtraje de cadenas

En general, el proceso de filtraje depende del dominio lingüístico y, ante todo, del tipo de la fuente lingüística: narrativa, artículo de periódico, diálogo por teclado,... De hecho, no poseemos ningún estudio sobre la influencia estadística de las diferentes fuentes lingüísticas.

El filtro, en realidad, utiliza una categorización de la entrada que es diferente de la que utiliza el proceso de parsing. De hecho es el analizador morfológico quien suministra al filtro la metalingüística de información morfológica necesaria. Merece señalarse que la ambigüedad morfológica ya ha sido reducida por los recursos del analizador (lexías, tiempos compuestos verbales, determinados conceptos del dominio,...).

La elección de técnicas de filtraje<sup>6</sup> fue decidida para impedir el excesivo backtracking del proceso de parsing. Tres técnicas se estudiaron como candidatas, el parsing determinista, el chart y el filtraje. La primera de ellas representa un difícil desafío para una gramática de propósito general y no se desestimó sino que, de hecho, el parser SP utiliza técnicas de 'previsión de futuro'<sup>7</sup> y una estrategia incremental. La segunda<sup>8</sup>, de hecho, fue desestimada ya que las herramientas no admitían un proceso eficiente de los necesarios 'assert' y 'retract' a bases de datos utilizadas por esta técnica del 'chart'. De esta forma el filtraje fue elegido como una opción básica aunque no necesariamente exclusiva de otras. De hecho, como el analizador morfológico se encarga de los tiempos compuestos,... y de las lexías y demás efectos léxicos no composicionales, tenemos ya de entrada una gran ventaja para reducir la más que posible dispersión estadística debida a estos fenómenos lingüísticos de distorsión de la fidelidad de las muestras. La dependencia de esta técnica respecto a la posibilidad de disponer de muestras de las fuentes analizar no supone problema, dado el destino que se le pretende dar dentro del

---

<sup>5</sup> La resolución de los cuatro tipos esenciales de errores tipográficos: letra extra, letra substituida, letra omitida y baile de letras, se integra de forma natural y sencilla en MORFEO

<sup>6</sup> La tesis doctoral de Kallas [1987] es la mejor exposición que he encontrado de estas técnicas markovianas

<sup>7</sup> El 'look-ahead', técnica que se halla ya implementada en PARSIFAL. Véase [Marcus 1980]

<sup>8</sup> Para los 'charts' véanse los artículos de M. Kay (1982) "Algorithm schemata and data structures in syntactic processing" En *Text Processing: Text Analysis and Generation, Text Typology and Attribution*, Proc. of the Symposium 51; y (1985) "Unification in Grammar" en: V. Dahl y P. Saint-Dizied eds. *Natural Language Understanding and Logic Programming*, Elsevier.

proyecto. Incluso muestras de pequeño tamaño parecen resultar rápidamente en datos probabilísticos de elevada fiabilidad.

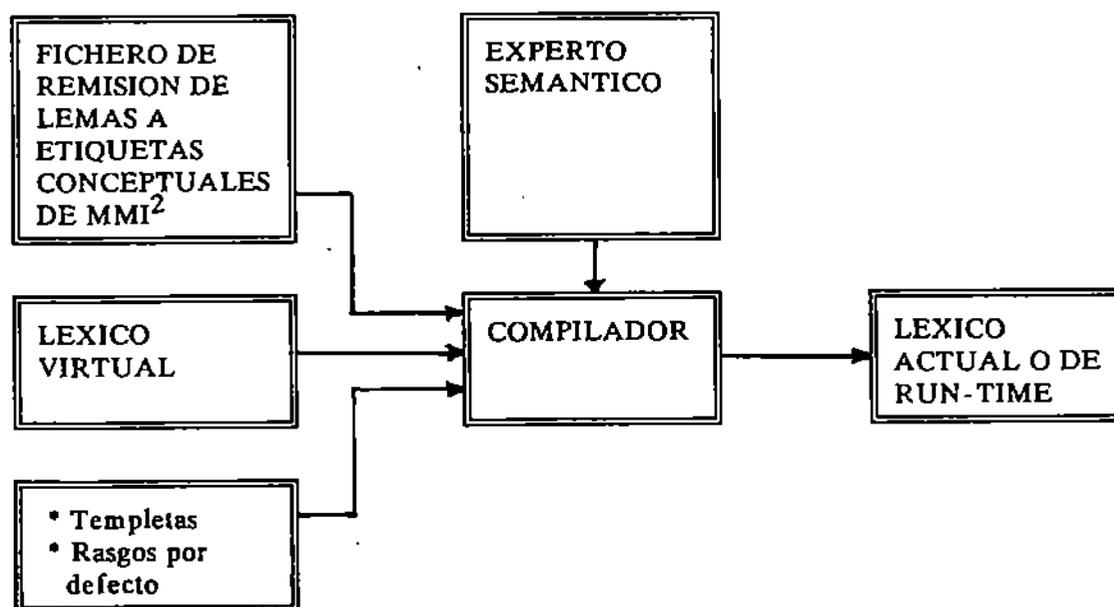
El filtro del SPM hace un ranking de las cadenas morfológicas alternativas para pasar al parser la cadena de posición más alta de probabilidad frente a la muestra de la fuente lingüística objetivo. El resto de cadenas son sucesivamente solicitadas por el parser, una a una, cuando aparece un fallo total en su respectiva inmediata anterior.

### 5.6. Los léxicos

La constitución del Léxico del SPM para una aplicación experta concreta implica una fase de compilación a partir de diferentes fuentes. El objetivo es gestionar el tamaño del conjunto de lemas para el run-time, así como las ambigüedades de acepción aceptadas para cada una de las piezas léxicas, para conseguir una eficiencia elevada en el run-time de las consultas. El Léxico está compilado conjuntamente, además de con otros ficheros operativos, con los de rasgos por defecto y macros, con la información sintáctica contenida en el Léxico virtual de la lengua española, con información contenida en el Experto Semántico que modela el conocimiento del mundo del interfaz MMI<sup>2</sup> durante sus etapas activas de conexión a una aplicación experta de análisis y diseño, y con un fichero de remisión de cada acepción semántica o lema del Español a las etiquetas de ese Experto Semántico que conjuntamente expresan los desplazamientos de su significado para el mundo de MMI<sup>2</sup> en un dominio de aplicación.

El objetivo principal de la compilación es expandir las entradas léxicas del Léxico Virtual por medio de plantillas y rasgos por defecto, seleccionar los lemas que remiten a etiquetas conceptuales del mundo actual de MMI<sup>2</sup>, aumentando, si cabe, las distinciones lemáticas o desplazamientos de significado del Léxico Virtual con las derivadas de ese mundo actual. El vínculo entre el Léxico y el Experto Semántico se da por medio del aspecto TYPE del 'frame' de los lemas. El procedimiento de compilación devuelve el Léxico actual para un entorno MMI<sup>2</sup>.

El esquema ilustra el proceso que construye el Léxico actual.



Las estructuras de información del Léxico contienen diferentes tipos de datos que representan toda la información pertinente sobre los lemas, para el propósito de la etapa de parsing.

Estos tipos de dato son:

- Rasgos convencionales de formato <ATRIBUTO> = <VALOR>.
- Rasgos booleanos.
- Templetas.

Los rasgos convencionales aceptan valores atómicos y disyuntivos.

### 5.7. El parser SP

El parser SP<sup>9</sup> está escrito en BIMproLog. Actualmente consta de un centenar de reglas de tipos diversos.

El parser procede en fases:

En la fase I, el resultado del análisis morfológico activa dos tipos de reglas: patrones (fechas, porcentajes, números telefónicos, etc.) y reglas combinatorias entre categorías terminales, y construye constituyentes mínimos aptos para ser reconocidos con criterios puramente deterministas<sup>10</sup>. Esta reducción tiene la ventaja de reducir también apreciablemente el espacio de búsqueda (y el subsiguiente 'backtracking') y, de esta forma, se puede concentrar el análisis en las relaciones de dependencia entre estos grupos.

La fase II combina sucesivamente los grupos elementales de derecha a izquierda, por medio de un procedimiento ascendente que incorpora además métodos de previsión ('look-ahead') capaces de abordar sólo los caminos de búsqueda que intentaría un parser de tipo descendente<sup>11</sup>.

El procedimiento debe obtener una representación funcional única para la frase.

El parser de MMI<sup>2</sup> es incremental en el sentido que el procesamiento sintáctico y el semántico tienen lugar al mismo tiempo.

Esta estrategia de análisis paralelo de la sintaxis y la semántica es muy eficiente (cf. Mellish 1983, Lytinen 1987) ya que hace innecesario construir representaciones sintácticas independientes en cualquier momento del proceso de análisis<sup>12</sup>.

Llevando este punto de vista a su conclusión lógica, el papel de la sintaxis se reduce a marcar con categorías las representaciones semánticas, y relacionarlas mediante reglas.

---

<sup>9</sup> Mayor detalle sobre el parser SP y sus reglas puede encontrarse en J.C. Ruiz " Un parser basado en la sintaxis de dependencias ", Procesamiento del Lenguaje Natural, Boletín nº 9, Marzo, 1991

<sup>10</sup> Francopoulo (1988) adopta una idea similar en su tesis sobre el aprendizaje automático de reglas gramaticales

<sup>11</sup> Se trata de una tabla de accesibilidad, idéntica a la usada en el parser BUP de Matsumoto et al. (1983, 1985)

<sup>12</sup> Existen bastantes pruebas de experimentos psicolingüísticos de que la comprensión humana se caracteriza por una estrategia semejante. Véase Pareschi (1987) para más referencias

### 5.8. El paso al formalismo CMR

Las representaciones funcionales se convierten en fórmulas de CMR, mediante un procedimiento que asigna cuantificadores y calcula su alcance respectivo<sup>13</sup>.

#### **Referencias**

- J. Allen (1983) "Recognizing Intentions from Natural Language Utterances". En: M.Brady & R.C. Berwick, **Computational Models of Discourse**. Cambridge Mass. The MIT Press.
- JL. Binot, P. Falzon, R. Pérez, B. Peroche, N. Sheehy, J. Rouault, M. Wilson (1991) **Architecture of a Multimodal Dialogue Interface for Knowledge-Based Systems**. Revista CEE, (a aparecer).
- S. Carberry (1990) **Plan Recognition in Natural Language Dialogue**. The MIT Press.
- D. Davidson (1980) **Essays on Actions and Events**. Clarendon Press, Oxford.
- G. Francopoulo (1988) **Un analyseur de français avec apprentissage inductif de la syntaxe**. Ph.D.Dissertation, Université Pierre et Marie Curie.
- H.P. Grice (1975) **Logic & Conversation**. En Cole & Morgan.
- N. Ghali, K. Mardia & N.Sheehy (1990) **An Algorithm to Capture Designer's Use of Shorthand 'Gesture'**. Report for the Second Technical Audit. University of Leeds.
- B.J. Grosz (1977) **The representation and Use of Focus in Dialogue Understanding**. Technical Note 151 SRI Project 5844.
- M.A.K. Halliday (1985) **An Introduction to Functional Grammar**. Edward Arnold ed.
- G. Hirst (1987) **Semantic Ambiguity and the Resolution of Ambiguity**. Cambridge University Press.
- J.R. Hobbs (1985) **Ontological promiscuity**. Proc. 23rd Annual Meeting of Association for Computational Linguistics, Chicago, June 1985, pp.61-69.
- G. Ioup (1975) "Some Universals for Quantifier Scoping". En J. Kimball, ed. **Syntax and Semantics Vol. 4**, Academic Press.
- R. Jackendoff (1977) **X' Syntax: A Study of Phrase Structure**. The MIT Press.
- R. Jackendoff (1983) **Semantics & Cognition**. The MIT Press.
- G. Kallas (1987) **Resolution des Solutions multiples en analyse morphologique automatique des langues naturelles. Utilisation des modèles de Markov**. Thèse tenue à l'Université des Sciences Sociales de Grenoble, CRISS.
- Krishnamurti (1988) **Representational Semantics for Graphical Discourse**. ESPRIT/ACORD task 3.6. deliverable. EdCAAD report, Jan. 1988.
- S. Lytinen (1987) "Integrating syntax and semantics" En S. Nirenburg (ed.) **Machine**

---

<sup>13</sup> Vid. Ioup [1975]

Translation, Cambridge University Press, pp. 302-316.

M.P.Marcus (1980) **A Theory of Syntactic Recognition for Natural Language**, MIT Press.

Y. Matsumoto, H. Tanaka, H. Hirakawa, and H. Yasukawa, H. (1983) "BUP: A Bottom-Up Parser Embedded in Prolog". En **New Generation Computing**, 1.

Ch. Mellish (1983) "Incremental semantic interpretation in a modular parsing system". En K. Sparck-Jones and Y. Wilks (eds.), **Automatic Natural Language Parsing**, Ellis Horwood, pp. 148-155.

Y. Matsumoto, M. Kiyono and H. Tanaka (1985) "Facilities of the BUP Parsing System". En V. Dahl y P. Saint-Dizier (eds.) **Natural Language Understanding and Logic Programming**, vol. I, North-Holland.

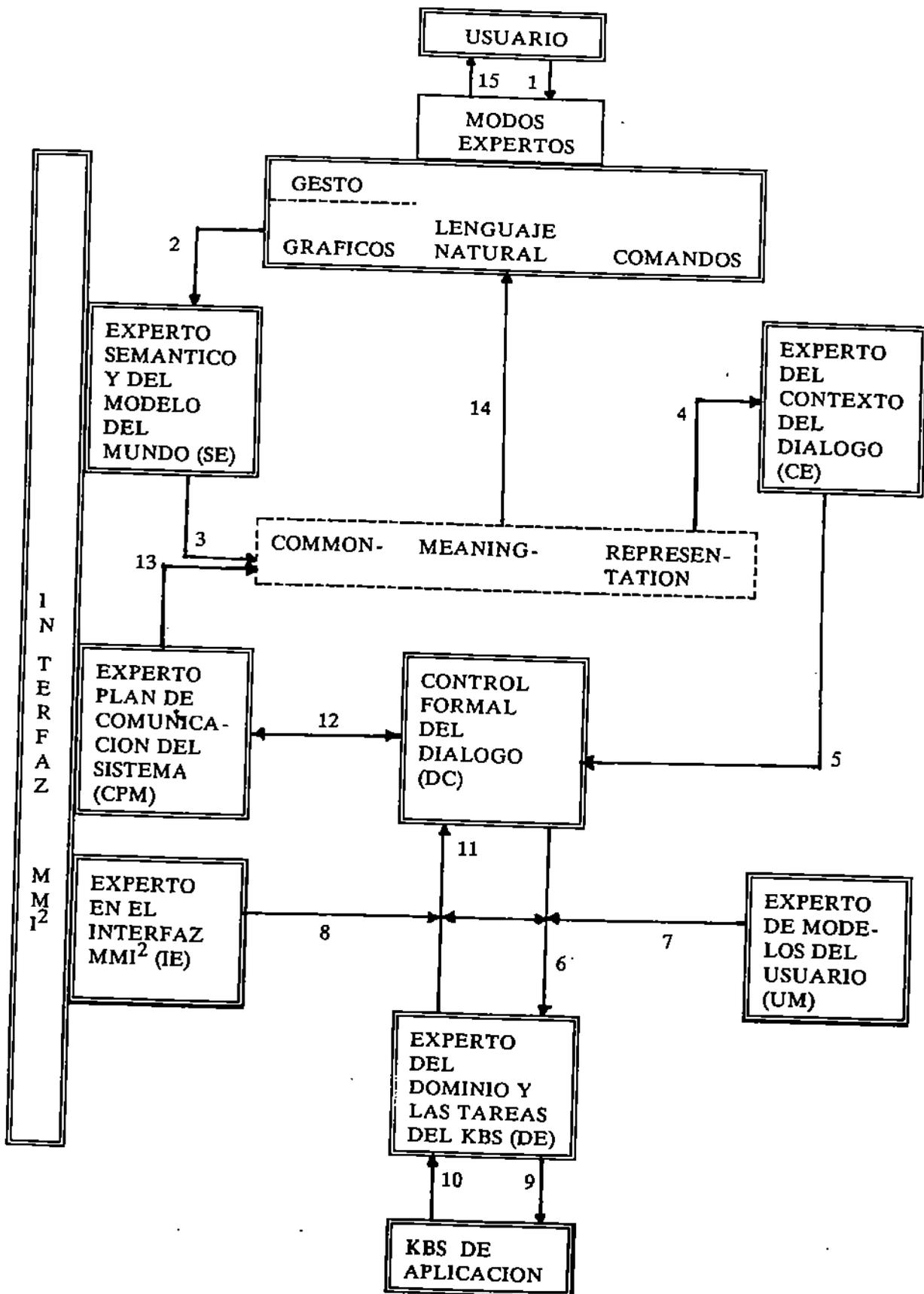
R. Pareschi (1987) "Combinatory Grammar, Logic Programming and Natural Language Processing". In N.Haddock, E. Klein and G. Morrill (eds.) **Categorial Grammar, Unification Grammar and Parsing**, Centre for Cognitive Science, University of Edinburgh, pp. 85-113.

W.V. Quine (1985) "Events and Reification". In **Actions and Events. Perspectives on the Philosophy of D. Davidson**, ed. por E. Lepore & B.M. McLaughlin, Basil Blackwell, Oxford, U.K.

C.L. Sidner (1983) "Focusing in the Comprehension of Definite Anaphora". En: M.Brady & R.C. Berwick, **Computational Models of Discourse**. Cambridge Mass. The MIT Press.

J.F. Sowa (1984) **Conceptual Structures**. Addison-Wesley Publ.

K. Sparck Jones (1987) **Realism about User Modelling**. Technical Report nº 111, University of Cambridge Computer Laboratory.



Esquema del Interfaz MMI<sup>2</sup>

Esquema de la arquitectura del SPM

