

Juan Falgueras, Antonio Luis Carrillo, Antonio Guevara

Departamento de Lenguajes y Ciencias de la Computación, Universidad de Málaga

<juanfc, carrillo, guevara}@cc.uma.es>

Resumen: la investigación y el desarrollo de técnicas de diseño e implementación de interfaces de usuario usables es todavía una amalgama de metodologías provenientes de otras tantas áreas de investigación. En este artículo presentamos algunos de los modelos y técnicas de diseño de interfaces de usuario más utilizados por los especialistas en éste área, categorizados en cuatro enfoques básicos (empírico, cognitivo, predictivo y antropomórfico), si bien veremos que estas divisiones no son nítidas ni rígidas, y que la mayoría de las técnicas incluyen características de más de una de ellas. La taxonomía presentada utiliza como criterio el objetivo buscado y la metodología empleada. Esta taxonomía no se ha planteado antes de una manera tan completa ni se han buscado otros criterios de división adecuados. Esto proporciona una visión en altura que sitúa y enriquece a cada categoría, permitiendo la identificación de la amplia cantidad de material existente en este terreno, y ayuda a avanzar en la búsqueda de nuevos enfoques a los nuevos problemas.

Palabras clave: interfaces de usuario, interacción persona-ordenador, modelos de diseño, modelo empírico, modelo cognitivo, modelo predictivo, modelo antropomórfico.

1. Introducción

La taxonomía propuesta no pretende realizar una clasificación ni de los estilos de interacción (interfaz por línea de comandos, menús y formularios, manipulación directa e interacción asistida), ni de los paradigmas de interacción (ordenador de sobremesa, realidad virtual, computación

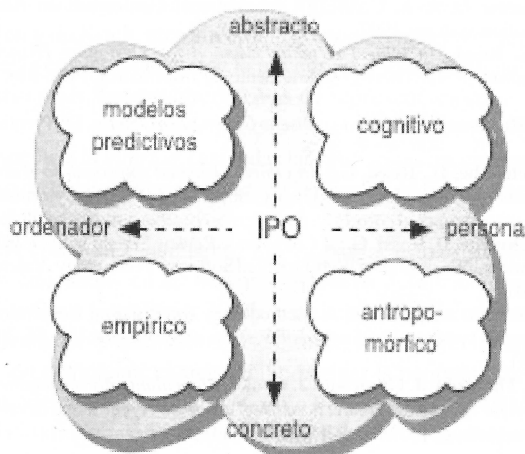
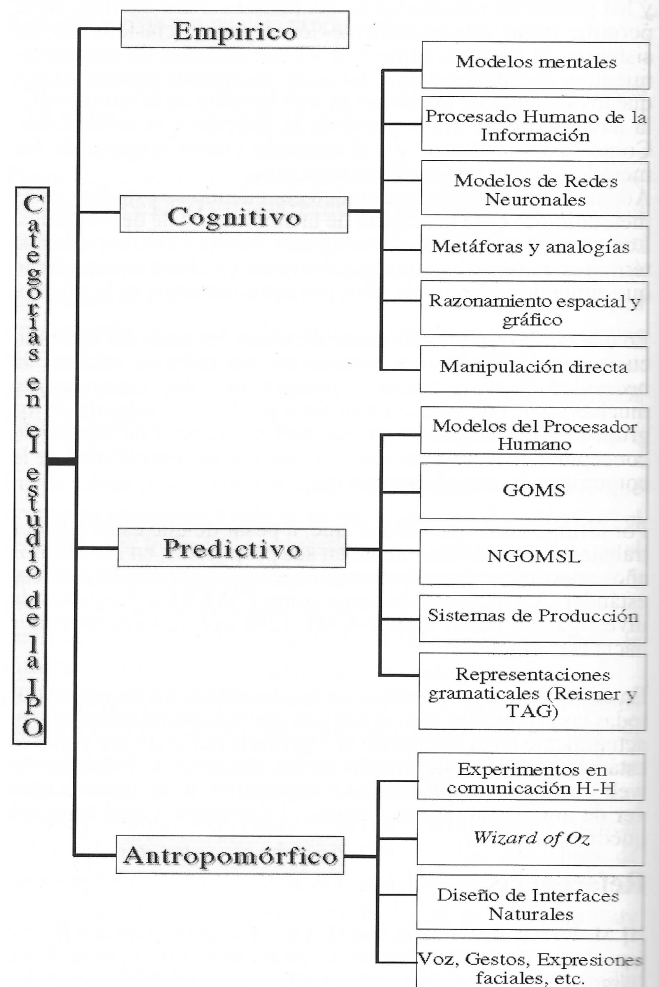


Figura 1. Categorías en el estudio de la Interacción Persona-Ordenador.

Enfoques en el estudio de las interfaces de usuario

ubicua y realidad aumentada) [1], ni siquiera de las técnicas formales de diseño y modelado de la interfaz de usuario [6,1]. Este trabajo se centra en el objetivo y método general usado. Un método puede querer entender el proceso de reconocimiento humano (cognitivo), o evaluar prácticamente tiempos de trabajo, pero estos métodos pueden aplicarse independientemente de tratarse de interfaces tipo líneas de comando, o incluso trabajar con la ergonomía de cualquier artefacto. La importancia de esta clasificación radica en que: 1) no existe ninguna y, 2) ayuda a entender la situación general actual del tema.

Desde las distintas áreas de investigación (sicología, informática, ergonomía, etc.) que se acercan al estudio de la interfaz



persona ordenador, se están aplicando teorías, metodologías y técnicas de diseño y evaluación substancialmente diferentes, persiguiendo objetivos muy distintos. Así, independientemente del área de partida, se pueden aplicar técnicas y estudios muy consolidados sobre los procesos de comunicación humana y de reconocimiento cognitivo, tomados ya bien de la visión antropomórfica, ya bien de la psicología cognitiva. Desde otros puntos de vista, más técnicos y actualmente en un desarrollo vertiginoso, la investigación se plantea en relación con mejoras en técnicas gráficas, velocidad, reconocimiento de la expresión humana, etc.

Según los trabajos de Eberts [4] y Falgueras [6], podemos categorizar en cuatro los modelos o enfoques al estudio de interfaces de usuario: *empírico*, *cognitivo*, *predictivo* y *antropomórfico* (figura 1). Aunque hay que tener en cuenta que estas divisiones no son necesariamente nítidas ni rígidas en todos los casos, ya que como veremos, un gran número de estudios podrían quedar encuadrados dentro de más de una de ellas. Por ejemplo, *muchas técnicas del modelo predictivo están basadas en teorías cognitivas, pudiéndose por tanto clasificar bajo ambos enfoques*. Por otro lado está la cuestión del modelo y las técnicas formales que se puedan emplear, pero estos no son más que herramientas para un objetivo. Los modelos predictivos fundamentalmente, pero también algunos cognitivos y antropomórficos, pueden ser vistos como desarrollos de formalizaciones con distintos orígenes, hecho que podría llevar hacia una *taxonomía basada en la formalización*, que no es la que buscamos. Ya que, a pesar del hecho de que algunas categorías de estudio dependan muy estrechamente de un modelo formal, también existen otras que no siguen siquiera un método de formalización o bien admiten cualquiera de ellos.

2. Modelo Cognitivo

Relacionado con el estudio de la persona, las teorías cognitivas (cuyas bases se encuentran en la psicología cognitiva), establecen cómo el ser humano percibe, guarda y recupera la información de su memoria a corto plazo, cómo manipula esta información para tomar decisiones y cómo da soluciones y resuelve problemas. Además, ven al usuario como un ser adaptativo, flexible e involucrado de manera activa en la interacción con el entorno en búsqueda de soluciones.

Este modelo se basa en la existencia de tres puntos de vista: (1) el *modelo conceptual*, o descripción del sistema desde el punto de vista del ingeniero que lo ha diseñado, por lo que se tratará de una descripción precisa y coherente. (2) El *modelo mental* o imagen que el usuario se hace sobre cómo es y cómo funciona el sistema, y que será el que ayude y guíe al usuario en su toma de decisiones. Y (3) la *interfaz*, que es el conjunto de estética, representaciones utilizadas, documentación, etcétera, que usa el diseñador para tratar de guiar al usuario hacia un modelo mental lo más cercano al conceptual y lo más práctico posible.

En general, desde el enfoque cognitivo, la interacción persona ordenador se concibe como la *presentación* de problemas a un usuario que tiene que *resolverlos*. Por esto muchas de las teorías que se aplican en psicología se aplican también aquí. Particular-

mente han resultado de gran influencia las teorías de Newell y Simon [11] sobre la **resolución de problemas**, que se puede considerar como una serie interconectada de objetivos y subobjetivos en jerarquía, lo que conduce a un modelo muy global del ser humano y relativamente sencillo de su comportamiento, y gracias a lo cual se han logrado derivar modelos formales como los de **Bovair y Polson** [2], **Kieras y Polson** [8], o **Card y otros** [3]. Son modelos cognitivos porque aportan mecanismos que permiten analizar y describir el conocimiento que el usuario posee o debe poseer del sistema. Con estos estudios se pretende entender y prever el comportamiento humano en la interacción.

Generalmente, la consideración de los aspectos cognitivos en la interfaz aumenta su calidad muy por encima de otras posibles mejoras técnicas o estéticas basadas en funcionalidades del software o del hardware. Por ejemplo, dado que la **analogía** y la **metáfora** son formas especialmente frecuentes de aprendizaje, un recurso importante tomado de la psicología cognitiva es la identificación de las tareas con las que interactuar, mediante metáforas dramatizadas [9]. Se trata pues, de presentar al usuario a través del diseño de la interfaz, alguna información familiar que reconozca fácilmente [1]. Un ejemplo importante y muy popularizado es el de la «*mesa de escritorio*», con la que se pretende reproducir una sobremesa de oficina y todos los objetos que tenemos en ella y en sus alrededores. Así, viendo la imagen de una «*papelera*», es de suponer que el usuario entienda que está relacionada con el hecho de deshacerse de objetos que ya no necesita; una «*carpeta*» habría de servir para guardar documentos, etc.

También está ampliamente demostrado que muchos problemas se resuelven mejor mediante la **manipulación directa** de imágenes, término acuñado por Ben Shneiderman en 1983 y que se refiere al hecho de que los usuarios puedan llevar a cabo las tareas y observar los resultados de forma inmediata, sirviéndose de la representación visual tanto de los objetos como de las acciones. Esto permite acciones rápidas, incrementales y reversibles que provocan un efecto visible inmediato en el objeto seleccionado, al mismo tiempo que se proporciona al usuario la sensación de que es *él* quien controla directamente los objetos representados en la pantalla y le permite alcanzar rápidamente un alto grado de familiaridad.

Aunque esta técnica pudiera parecer idónea en todos los casos, no es así, uno de los inconvenientes de la manipulación directa es que no todas las tareas involucran objetos reconocibles gráficamente y no todas las acciones se pueden hacer directamente.

Por otro lado, el hecho de que se requiera la atención del ser humano para que este pueda procesar la información hace que otro de los aspectos y problemas más importantes contemplados en este área de estudio sean los de los «*focos de atención*» y el de los «*límites de la atención*».

Quizás, uno de los principales *inconvenientes* del modelo cognitivo puede ser la ambigüedad de los términos utilizados, que no sugieren al diseñador formas concretas de realización.

3. Modelo Predictivo

Es un modelo eminentemente práctico, en el cual se da un enfoque básicamente (pero no únicamente) ingenieril al proceso de diseño de la interfaz de usuario, generalmente mediante técnicas formales, y en el que se tratará de predecir la eficiencia de la interacción entre el humano y el ordenador. Es una medida de viabilidad antes del prototipado, y por lo tanto, básicamente teórica y basada en modelos abstractos. Además, varias de estas técnicas permiten en mayor o menor grado, tanto plantear formalmente la interfaz que se desea construir y describir su funcionamiento (función generativa), como analizarla y evaluarla con objeto de poder depurarla y mejorarla.

El diseñador deberá por tanto, realizar tareas de modelado, considerar posibles diseños, modelar las tareas de estos diseños y elegir uno de ellos, el mejor diseño, basándose en predicciones cuantitativas y estimaciones precisas de tiempos para las tareas. Se basan en el análisis y por lo tanto tienen el *inconveniente* de requerir la interpretación de las tareas.

El **modelo de procesador humano** (MHP) desarrollado por Card, Moran y Newell [3], al igual que la mayoría de teorías de procesamiento de la información, parte de la hipótesis de que la información se procesa por etapas, concretamente en tres: la perceptual, la cognitiva y la motora. Así, mediante el análisis de la información a procesar y de los ciclos de tiempo que requiere cada etapa, es posible hacer predicciones temporales y predecir también errores de uso.

De entre las técnicas y modelos de ingeniería existentes para el diseño y evaluación de interfaces de usuario, quizás los más extendidos en la práctica son **los modelos GOMS**, inicialmente propuestos por Card, Moran y Newell [3] y basados en anteriores trabajos de Newell y Simon [11], relacionados con la teoría cognitiva. Los cuatro más conocidos son la formulación original CMN-GOMS, el NGOSML, CPM-GOMS y el KLM (que es el que desciende a un mayor nivel de detalle). GOMS es el acrónimo de Goals, Operators, Methods, and Selection rules, componentes que se utilizan como elementos de construcción en los modelos GOMS. Los «Objetivos» (*Goals*) se refieren a lo que el usuario está intentando conseguir (por ejemplo, cerrar una ventana de la interfaz), usualmente especificado de manera jerárquica. Los «Operadores» son el conjunto de operaciones atómicas con las cuales se compone la solución deseada, y que pueden afectar realmente al sistema (por ejemplo, pulsar una tecla de control) o tan sólo al estado mental del usuario (leer un cuadro de texto). Los *Métodos* representan entonces secuencias de operadores, agrupados juntos para satisfacer un objetivo simple, pudiendo existir diversos caminos para conseguir el mismo objetivo, y cada uno de los cuales constituirá un método. Por ejemplo, para cerrar una ventana, se puede pulsar el botón de cierre (método-Botón), pulsar alguna combinación de teclas (método-Comando) o seleccionar un elemento de menú (método-Menú). Las *reglas de Selección* permitirán decidir o escoger el método a utilizar para alcanzar un objetivo concreto, si se da el caso de poder elegir entre varios, como en el ejemplo anterior. Así, si se

trata de un usuario experto en el uso de la aplicación, escogerá el método-Comando para cerrarla, y en caso contrario, el método-Menú.

La diferencia entre un objetivo y un operador depende simplemente del nivel de detalle escogido por el analista. Para satisfacer un objetivo, el analista escoge un método que usa operadores que especifican los detalles del cómo se satisface. Por el contrario, los operadores no son susceptibles de posteriores descomposiciones.

Se podría decir que GOMS, más que un modelo, que lo sería el de Procesador-Humano, es una técnica de análisis de tareas en términos de reglas de producción. Estas reglas de producción pueden ser después explotadas de diversas formas, a nivel de evaluación de tiempos, de estimación de complejidad, de requisitos de diseño, etc.

El método NGOSML (*Natural GOMS Language*) formulado por Kieras [7], elimina algunas de las ambigüedades del resto de métodos de esta familia GOMS, dota al lenguaje de más naturalidad y permite completar las estimaciones o predicciones cuantitativas (de tiempos), con otras de tipo cualitativo. Además, existe una buena documentación sobre él y métodos prácticos de construcción. Podemos hacernos una idea mediante el siguiente ejemplo:

```

...
Método para el objetivo: Cortar-texto.
  Paso 1. Acometer objetivo: Resaltar-texto.
  Paso 2. Retener que el comando es CUT, y acometer
objetivo:
      Activar-comando.
  Paso 3. Volver con objetivo cumplido.
Regla de selección para el objetivo: Resaltar-
texto.
  Si el texto es una palabra, entonces acometer
objetivo:
      Resaltar-palabra.
  Si el texto es arbitrario, entonces acometer
objetivo:
      Resaltar-texto-arbitrario.
Método para el objetivo: Resaltar-palabra.
  Paso 1. Determinar una posición dentro de la
palabra
  Paso 2. Mover el cursor a la mitad de la palabra.
  Paso 3. Hacer Doble-click con el botón del ratón.
  Paso 4. Verificar que el texto correcto ha sido se-
leccionado.
  Paso 5. Volver con objetivo cumplido.
...

```

Los **sistemas de reglas de producción** han sido utilizados para modelar los procesos cognitivos humanos desde los estudios de Anderson en 1976. Esta técnica, proveniente del área de la Inteligencia Artificial, permite almacenar conocimientos acerca de una entidad para posteriormente procesar esta información para resolver problemas o tomar decisiones basadas en las condiciones del entorno. Las reglas son de tipo «si-entonces». Los objetivos, operadores, métodos y reglas de los modelos GOMS pueden convertirse en reglas «si-entonces» utilizables para modelar el comportamiento de la interfaz.

Dado que las gramáticas cognitivas se desarrollaron para poder describir las reglas de uso de los lenguajes, y como los programas interactivos se pueden conceptualizar en térmi-

nos de lenguajes interactivos, también se puede utilizar **gramáticas** cognitivas para describir las tareas de la interacción persona ordenador.

En el enfoque que hace Reisner [13], se utilizan los formularios de Backus-Naur (BNF) para describir la gramática de un diálogo, presentándolo a un nivel meramente sintáctico e ignorando el significado de las acciones. Las descripciones BNF describen las acciones del usuario, no la percepción del mismo de las respuestas del ordenador. En el siguiente ejemplo:

```
...
BorrarFichero ::= Seleccionar + ArrastrarAPapelera
Seleccionar ::= (RecordarInformacionSeleccionar)
+ HacerSeleccionar
...
```

se puede observar que las acciones propiamente cognitivas se encierran entre paréntesis, y las motoras observables no. Estas descripciones pueden ser analizadas de varias formas. Por ejemplo, mientras más reglas requiera una interfaz, más complicada será, aunque esta medida es incluso sensible al estilo empleado al describir las propias reglas. De hecho no ha tenido muchos seguidores y se plantea que sencillamente, los **diagramas de transición de estados** son mejores.

El método TAG (*Task Action Grammar*) de Payne y Green [12], utiliza también descripciones gramaticales BNF. Esta notación jerarquiza rápidamente conceptos y acciones finales evidenciando la estructura y jerarquía del proceso interactivo. Los símbolos no terminales son *Goals* y las acciones elementales son símbolos terminales. Este método facilita la medición de la *consistencia* de la interfaz, la aparición de patrones repetidos (a mayor número de ellos, mayor consistencia y mayor facilidad de aprendizaje). Sin embargo, aún no se han desarrollado estudios que aporten medidas ni den idea de la mayor o menor eficiencia de la interfaz, y las descripciones son largas y difíciles de leer.

A modo de resumen, podemos concluir que los modelos gramaticales son especialmente útiles para la evaluación de la consistencia del diseño de la interfaz y para sistemas complejos y con pocos detalles de tipo cognitivo de bajo nivel o donde estos no sean esenciales. En otras situaciones suelen ser más apropiadas otras técnicas, basadas por ejemplo en el MHP, como el NGOSML. Por otro lado, respecto a tiempos de ejecución, sólo los modelos MHP, los de la familia GOMS y los sistemas de producción pueden ser utilizados para hacer predicciones cuantitativas.

4. Modelo Empírico

El diseño se basa en la propia experiencia del diseñador y en la recogida de *principios* generales de interacción, *recomendaciones* propias del sistema de trabajo y en *estándares* oficiales [1]. Los resultados tienen que verse respaldados posteriormente mediante tests de usabilidad.

En general, la interacción persona-ordenador depende de tantos factores **no formalizables ni previsible**s, que se necesitarán comprobaciones sobre situaciones reales, con

interfaces de usuario concretas ya diseñadas (en prototipos) o aplicaciones finales. El diseñador deberá recoger y analizar los resultados de estudio de campo en los que se observan comportamientos y se toman medidas durante el uso de elementos particulares de la interacción persona-ordenador (Moher y Schneider [10] y Embley [5]).

La principal *ventaja* de este modelo es que ofrece una alternativa a la intuición en la determinación del mejor diseño. De hecho, algunos experimentos de campo han invalidado teorías altamente intuitivas como la de la ventaja de las interfaces gráficas en determinadas áreas, donde se ha encontrado experimentalmente una mayor adecuación de las interfaces textuales. En otros casos, la intuición se puede ver corroborada por el experimento.

Sin embargo, presenta *aspectos negativos* importantes. La experimentación aislada no es adecuada en muchos casos, se pueden diseñar experimentos sin una investigación previa en las bases conductuales de los usuarios, a veces los usuarios sometidos a evaluación no son los adecuados (como por ejemplo, al emplean sujetos inexpertos), lo que lleva sencillamente a hacer imposible la generalización de los resultados. En muchos casos, el modelo experimental viene a surgir de carencias de orientación teórica y no induce a crear ninguna línea nueva. La mayoría de los experimentos se planifican para resolver problemas particulares en momentos concretos del desarrollo. Sólo puede ser utilizado para comprobar la bondad o falta de adecuación de una nueva interfaz, de acuerdo a sus expectativas de accesibilidad o eficiencia.

5. Modelo Antropomórfico

La idea básica en este modelo es tratar de alcanzar una comunicación fluida con la máquina siguiendo el proceso natural de comunicación persona-persona (familiar e intuitivo), intentando dotar a la interfaz con cualidades y gestos humanos reconocibles, como podrían ser el lenguaje natural, reconocimiento de imagen, texto, movimientos, etc.

Este modelo destaca la idea del «trato amigable con el usuario» y la «naturalidad» de la interfaz. Se dice que la interfaz es amigable si el sistema reacciona ante el usuario de una manera «comprensiva», evitando los mensajes crípticos, poniéndose en su lugar y ayudándole a trabajar cómodamente, creando un ambiente de concordia en el que los mutuos errores se acepten y corrijan de la manera más cómoda para ambos. «Natural» se refiere a que la persona y el ordenador actúen como compañeros, a que se equilibre la creatividad y a que interactúen de manera grata.

En general, se utilizan tres técnicas en este modelo. En la primera, **imitación**, la interacción real humano-humano se utiliza, después de una meticulosa observación, como modelo y medio de recogida de información vital para las decisiones en el diseño. Esta interacción se estudia o bien en un complejo, controlado y costoso laboratorio de observación, o bien en un menos controlado ambiente no ya de laboratorio, utilizando los datos posteriormente en el diseño de la interfaz del sistema en estudio.

El segundo método es más intuitivo y menos costoso. En él se hace un **estudio comparativo** entre la interacción persona-ordenador que se produzca y la humano-humano, en busca de aspectos desnaturalizados y difíciles para una persona, planteando entonces, para los procesos más críticos, métodos de comunicación mediante la voz, lenguaje natural, etc.

El tercer método se basa en la propia **experiencia** de los investigadores o desarrolladores y en la aplicación de soluciones ya existentes en anteriores aplicaciones, lo que puede representar para el usuario el encuentro con nuevas formas de ver los procesos ya conocidos, simplificando incluso los de comunicación entre personas mediante un argot o jerga familiar al usuario del sistema.

Entre las principales *desventajas* del modelo antropomórfico se encuentra el hecho de que el uso precipitado de analogías en la comunicación limita la incorporación de nuevas habilidades, que en una interacción hecha de familiaridad, no serían reconocidas. El desarrollo de la interacción antropomórfica dependerá en gran medida de los avances tecnológicos y en la mejora de las técnicas de inteligencia artificial. Por otro lado, algunos usuarios, especialmente los más expertos, pueden rechazar sistemas que les resulten excesivamente amigables o dialogantes.

6. Conclusiones

En este artículo hemos comparado algunos de los principales modelos y técnicas de diseño de interfaces de usuario utilizados por los especialistas en este área, categorizados en cuatro enfoques básicos, basándonos principalmente en la técnica y el objetivo general buscado por el investigador.

Los enfoques más utilizados en el proceso de análisis y diseño son el predictivo y el cognitivo. El primero, más ingenieril, pero basado en los estudios de los otros enfoques, ofrece técnicas, generalmente formales, que permiten en mayor o menor grado, tanto modelar la interfaz como realizar su evaluación, al ofrecer la posibilidad de realizar predicciones cuantitativas (y en algunos casos, también cualitativas).

El modelo cognitivo se centra más en aspectos de tipo percepción, almacenamiento y procesado de la información, y en el cómo se dan soluciones y se resuelven problemas. Tiene sus bases en la psicología cognitiva. Sugiere la realización de metáforas y analogías, manipulación directa de imágenes, captación de la atención del usuario, etc. Sustenta conceptualmente a los modelos predictivos. Sin embargo, por lo general, no ofrece formas concretas de desarrollo.

Por otro lado, el enfoque antropomórfico persigue mejorar la interacción o comunicación persona-ordenador e intenta imitar el proceso natural de comunicación persona-persona. Destaca conceptos como «naturalidad» y «amigabilidad con el usuario». Sin embargo, tiene el inconveniente de depender en gran medida de los avances tecnológicos y de las técnicas de inteligencia artificial.

Por último, el enfoque empírico, más que ofrecer técnicas generativas, constituye una alternativa a la intuición en la

determinación del mejor diseño, centrándose en comprobaciones sobre situaciones reales, con interfaces de usuario ya diseñadas.

Al abordar el estudio de un sistema interactivo se deben considerar todos estos enfoques. Sus orígenes, sin embargo, son muy distintos, como hemos visto.

Referencias

- [1] Abascal, J. y otros; La Interacción persona ordenador. Libro electrónico editado por AIPO. Versiones disponibles en la red (<http://griho.udl.es/ipo/>) y en CD-ROM (ISBN Versión CD-ROM 84-607-2255-4), 2002.
- [2] Bovair, S. y Polson, D.; «The Acquisition and Performance of Text-Editing Skill: A Cognitive Complexity Analysis». *Human-Computer Interaction*, 5(1), 1-48, 1990.
- [3] Card, S.K., Moran T.P. y Newell, A.; *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum Associates. 1983.
- [4] Eberts, R.E.; *User Interface Design*. Prentice-Hall. 1994.
- [5] Embley, D.W.; «Empirical and Formal Language Design Applied to a Unified Control Construct for Interactive Computing». *International Journal of Man-Machine Studies*, 10(2), 197-216. 1978.
- [6] Falgueras, J.; Usabilidad informática; Diseño actual de la interfaz de usuario. Tesis doctoral. Departamento de Lenguajes y Ciencias de la Computación. Universidad de Málaga. 2000.
- [7] Kieras, D.E.; «Towards a Practical GOMS Model Methodology for User Interface Design». En M. Helander (Ed.), (Págs. 135-157). North-Holland, *Elsevier Science*. 1988. (ISBN 0444705368).
- [8] Kieras, D.E. y Polson, P.G. [1985]. «A Quantitative Model of the Learning and Performance of Text Editing Knowledge». En *Proceeding of ACM CHI '85 Conference on Human Factors in Computing Systems* (207-212). 1985
- [9] Laurel, B.; *Computers as Theatre*. Addison-Wesley. Reading-MA. 1991.
- [10] Moher, T. y Schneider, G.M.; «Methodology and Experimental Research in Software Engineering». *International Journal of Man-Machine Studies*, 16(1), 65-87. 1982.
- [11] Newell, A. y Simon, H.; *Human problem solving*. Prentice-Hall. 1972.
- [12] Payne, S.J. y Green, T.; «Task-Actions Grammars: a model of the mental representation of task languages». *Human-Computer Interaction*, 2(2), 93-133. 1986.
- [13] Reisner, P.; «Formal grammars and human factors design of an interactive graphics system». *IEEE Transaction on Software Engineering*. SE 7,2, 229-240. 1981.