

FUNDAMENTOS DE PROGRAMACIÓN

Algoritmos, estructura
de datos y objetos

Cuarta edición

Luis Joyanes Aguilar

Catedrático de Lenguajes y Sistemas Informáticos
*Facultad de Informática, Escuela Universitaria de Informática
Universidad Pontificia de Salamanca campus de Madrid*



MADRID • BOGOTÁ • BUENOS AIRES • CARACAS • GUATEMALA • LISBOA • MÉXICO
NUEVA YORK • PANAMÁ • SAN JUAN • SANTIAGO • SÃO PAULO
AUCKLAND • HAMBURGO • LONDRES • MILÁN • MONTREAL • NUEVA DELHI • PARÍS
SAN FRANCISCO • SIDNEY • SINGAPUR • ST LOUIS • TOKIO • TORONTO

PARTE I

Algoritmos y herramientas de programación

CONTENIDO

Capítulo 1. Introducción a las computadoras y los lenguajes de programación

Capítulo 2. Metodología de la programación y desarrollo de software

Capítulo 3. Estructura general de un programa

Capítulo 4. Flujo de control I: Estructuras selectivas

Capítulo 5. Flujo de control II: Estructuras repetitivas

Capítulo 6. Subprogramas (subalgoritmos): Funciones

Introducción a las computadoras y los lenguajes de programación

- 1.1. ¿Qué es una computadora?
 - 1.2. Organización física de una computadora
 - 1.3. Representación de la información en las computadoras
 - 1.4. Codificación de la información
 - 1.5. Dispositivos de almacenamiento secundario (almacenamiento masivo)
 - 1.6. Conectores de dispositivos de E/S
 - 1.7. Redes, Web y Web 2.0
 - 1.8. El *software* (los programas)
 - 1.9. Lenguajes de programación
 - 1.10. Breve historia de los lenguajes de programación
- RESUMEN

INTRODUCCIÓN

Las computadoras (ordenadores) electrónicas modernas son uno de los productos más importantes del siglo XXI ya que se han convertido en un dispositivo esencial en la vida diaria de las personas, como un electrodoméstico más del hogar o de la oficina y han cambiado el modo de vivir y de hacer negocios. Constituyen una herramienta esencial en muchas áreas: empresa, industria, gobierno, ciencia, educación..., en realidad en casi todos los campos de nuestras vidas. Son infinitas las aplicaciones que se pueden realizar con ellas: consultar el saldo de una cuenta corriente, retirar dinero de un banco, enviar o recibir mensajes por teléfonos celulares (móviles) que a su vez están conectados a potentes computadoras, escribir documentos, navegar por Internet, enviar y recibir correos electrónicos (e-mail), etc.

El papel de los programas de computadoras es fundamental; sin una lista de instrucciones a seguir, la computadora es virtualmente inútil. Los lenguajes de programación nos permiten escribir esos programas y por consiguiente comunicarnos con las computadoras. La principal razón para que las personas aprendan lenguajes y técnicas de programación es *utilizar la computadora como una herramienta para resolver problemas*.

En el capítulo se introducen conceptos importantes tales como la organización de una computadora, el *hardware*, el *software* y sus componentes, y se introducen los lenguajes de programación más populares C, C++, Java o C#.

1.1. ¿QUÉ ES UNA COMPUTADORA?

Las computadoras se construyen y se incluyen en todo tipo de dispositivos: automóviles (coches/carros), aviones, trenes, relojes, televisiones... A su vez estas máquinas pueden enviar, recibir, almacenar, procesar y visualizar información de todo tipo: números, texto, imágenes, gráficos, sonidos, etc. Estas potentes máquinas son dispositivos que realizan cálculos a velocidades increíbles (millones de operaciones de las computadoras personales hasta cientos de millones de operaciones de las supercomputadoras). La ejecución de una tarea determinada requiere una lista de instrucciones o un programa. Los programas se escriben normalmente en un lenguaje de programación específico, tal como C, para que pueda ser comprendido por la computadora.

Una **computadora**¹ es un dispositivo electrónico, utilizado para procesar información y obtener resultados, capaz de ejecutar cálculos y tomar decisiones a velocidades millones o cientos de millones más rápidas que puedan hacerlo los seres humanos. En el sentido más simple una computadora es “un dispositivo” para realizar cálculos o computar. El término sistema de computadora o simplemente computadora se utiliza para enfatizar que, en realidad, son dos partes distintas: *hardware* y *software*. El *hardware* es la computadora en sí misma. El *software* es el conjunto de programas que indican a la computadora las tareas que debe realizar. Las computadoras procesan datos bajo el control de un conjunto de instrucciones denominadas programas de computadora. Estos programas controlan y dirigen a la computadora para que realice un conjunto de acciones (instrucciones) especificadas por personas especializadas, llamadas *programadoras de computadoras*.

Los datos y la información se pueden introducir en la computadora por una **entrada** (*input*) y a continuación se procesan para producir una **salida** (*output*, resultados), como se observa en la Figura 1.1. La computadora se puede considerar como una unidad en la que se colocan ciertos datos (*entrada de datos*), se procesan y se produce un resultado (*datos de salida o información*). Los datos de entrada y los datos de salida pueden ser, realmente, de cualquier tipo: texto, dibujos, sonido, imágenes... El sistema más sencillo para comunicarse una persona con la computadora es mediante un teclado, una pantalla (monitor) y un ratón (*mouse*). Hoy día existen otros dispositivos muy populares tales como escáneres, micrófonos, altavoces, cámaras de vídeo, teléfonos inteligentes, agendas PDA, reproductores de música MP3, iPod, etc.; de igual manera, a través de *módems*, es posible conectar su computadora con otras computadoras a través de la red **Internet**.

Como se ha dicho antes, los componentes físicos que constituyen la computadora, junto con los dispositivos que realizan las tareas de entrada y salida, se conocen con el término *hardware* o sistema físico. El programa se encuentra almacenado en su memoria; a la persona que escribe programas se llama *programador* y al conjunto de programas escritos para una computadora se llama *software*. Este libro se dedicará casi exclusivamente al *software*, pero se hará una breve revisión del *hardware* como recordatorio o introducción según sean los conocimientos del lector en esta materia.

Una computadora consta de varios dispositivos (tales como teclado, pantalla, “ratón” (*mouse*), discos duros, memorias, escáner, DVD, CD, memorias *flash*, unidades de proceso, impresoras, etc.) que son conocidos como *hardware*. Los programas de computadora que se ejecutan o “corren” (*run*) sobre una máquina se conocen como *software*. El coste del *hardware* se ha reducido drásticamente en los últimos años y sigue reduciéndose al menos en términos de relación precio/prestaciones, ya que por el mismo precio es posible encontrar equipos de computadoras con unas prestaciones casi el doble de las que se conseguían hace tan sólo dos o tres años por un coste similar². Afortunadamente, el precio del *software* estándar también se ha reducido drásticamente, pero por suerte cada día se requieren más aplicaciones específicas y los programadores profesionales cada día tienen ante sí grandes retos y oportunidades, de modo que los esfuerzos y costes que requieren los desarrollos modernos suelen tener compensaciones económicas para sus autores.

¹ En España está muy extendido el término **ordenador** para referirse a la traducción de la palabra inglesa *computer*. El DRAE (Diccionario de la Real Academia Española, realizado por la Academia Española y todas las Academias de la Lengua de Latinoamérica, África y Asia) acepta, indistintamente, los términos sinónimos: computador, computadora y ordenador. Entre las diferentes acepciones define la computadora electrónica como: “máquina electrónica, analógica o digital, dotada de una memoria de gran capacidad y de métodos de tratamiento de la información capaz de resolver problemas matemáticos y lógicos mediante la utilización automática de programas informáticos”. En el *Diccionario panhispánico de dudas* (Madrid: RAE, 2005, p. 157), editado también por la Real Academia Española y la Asociación de Academias de la Lengua Española, se señala que el término **computadora** (del término inglés *computer*) se utiliza en la mayoría de los países de América, mientras que el masculino computador es de uso mayoritario en Chile y Colombia; en España se usa preferentemente el término ordenador, tomado del francés *ordinateur*. En este reciente diccionario la definición de computador es “Máquina electrónica capaz de realizar un tratamiento automático de la información y de resolver con gran rapidez problemas matemáticos y lógicos mediante programas informáticos”.

² A título meramente comparativo resaltar que el primer PC que tuvo el autor de esta obra, comprado en la segunda mitad de los ochenta, costó unos 5-6.000\$ y sólo contemplaba una unidad central de 512 KB, disco duro de 10 MB y una impresora matricial.

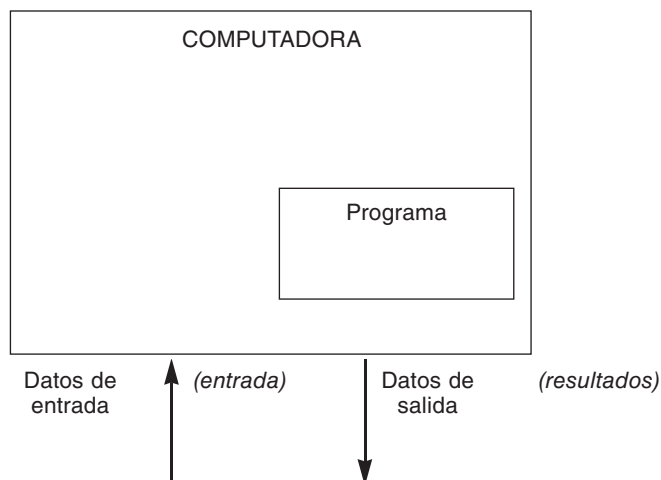


Figura 1.1. Proceso de información en una computadora.

1.1.1. Origen de las computadoras

La primera computadora digital que reseña la historia de la informática, se puede considerar, fue diseñada a finales de la década de los treinta por el Dr. John Atanasoff y el estudiante de postgrado Clifford Berry³ en la Universidad de Iowa (Iowa State University). Diseñaron la computadora para realizar cálculos matemáticos en física nuclear.

Sin embargo, la primera computadora electrónica digital de aplicaciones o propósito general se llamaba ENIAC y se terminó en 1946 en la Universidad de Pennsylvania, fue financiada por el Ejército de EE.UU. (U.S. Army). La ENIAC pesaba 30 toneladas y ocupaba un espacio de 30 por 50 pies. Se utilizaba esencialmente para predicciones de tiempo, cálculos de tablas balísticas, cálculos de energía atómica. Sus diseñadores fueron J. Prespert Eckert y John Mauchley.

En el mismo año de 1946, el Dr. John Von Neumann de Princeton University propuso el concepto de *computadora con programa almacenado* que consistía en un programa cuyas instrucciones se almacenaban en la memoria de la computadora.

Von Neumann descubrió que era posible que los programas se almacenaran en la memoria de la computadora y que se podrían cambiar más fácilmente que las complejas conexiones de cables y fijaciones de interruptores del ENIAC. Von Neumann diseñó una computadora basada en esta idea. Su diseño ha constituido el nacimiento de la computación moderna y ha dado origen a la denominada **arquitectura de Von Neumann** que es la base de las computadoras digitales actuales.

Estas computadoras primitivas utilizaban tubos de vacío como componentes electrónicos básicos. No sólo eran muy voluminosas, sino lentas y difíciles de manipular a la par que requerían usos y cuidados especiales. Los avances tecnológicos en semiconductores, transistores y circuitos integrados concluyeron en diseñar y fabricar las nuevas generaciones de computadoras que conducían a máquinas más pequeñas, más rápidas y más económicas que sus predecesoras.

En la década de los setenta, los fabricantes Altair (suele considerarse la primera microcomputadora de la historia) y Apple fabrican la primera microcomputadora de la historia. Steve Jobs y Stephen Wozniac construyen el Apple, la primera computadora doméstica de la historia. Por aquella época otras compañías que fabricaron microcomputadoras fueron Commodore, Radio Shack, Heathkit y en Europa, Sinclair que fabricó el mítico ZX Spectrum con el que aprendieron a programar y a jugar con videojuegos muchos de los grandes ingenieros, catedráticos, etc., de esta década. Eran computadoras que en aquella época no eran aceptadas por la comunidad profesional, las empresas y las industrias.

El 12 de agosto de 1981 IBM presentó en Nueva York y en otras ciudades norteamericanas, la primera computadora de escritorio de la historia, denominada por su inventor, IBM PC (Personal Computer, computadora personal de IBM), cuyo *software* fundamental fue desarrollado por una joven compañía conocida como Microsoft. El PC se convirtió en un éxito instantáneo hasta llegar a convertirse en un aparato o dispositivo electrónico⁴ de uso general, al

³ En su honor se conoce como computadora de Atanasoff-Berry.

⁴ *Coomodity*, el término por el que se conoce en inglés un dispositivo electrónico de consumo que se puede comprar en un gran almacén.

estilo de una TV o un equipo de música. Sin embargo, conviene recordar que el PC, tal como se le conoce en la actualidad, no fue la primera computadora personal ya que le precedieron otras máquinas con microprocesadores de 8 bits, muy populares en su tiempo, tales como *Apple II*, *Pet CBM*, *Atari*, *TRS-80*, etc., y el mítico *ZX Spectrum*, de los diferentes fabricantes citados en el párrafo anterior.

El término PC se utiliza indistintamente con el término genérico de *computadora de escritorio* o *computadora portátil* (*desktop*) o (*laptop*)⁵.

1.1.2. Clasificación de las computadoras

Las computadoras modernas se pueden clasificar en *computadoras personales*, *servidores*, *minicomputadoras*, *grandes computadoras* (*mainframes*) y *supercomputadoras*.

Las **computadoras personales (PC)** son las más populares y abarcan desde computadoras portátiles (*laptops* o *notebooks*, en inglés) hasta computadoras de escritorio (*desktop*) que se suelen utilizar como herramientas en los puestos de trabajo, en oficinas, laboratorios de enseñanza e investigación, empresas, etc. Los **servidores** son computadoras personales profesionales y de gran potencia que se utilizan para gestionar y administrar las redes internas de las empresas o departamentos y muy especialmente para administrar sitios Web de Internet. Las computadoras tipo servidor son optimizadas específicamente para soportar una red de computadoras, facilitar a los usuarios la compartición de archivos, de *software* o de periféricos como impresoras y otros recursos de red. Los servidores tienen memorias grandes, altas capacidades de memoria en disco e incluso unidades de almacenamiento masivo como unidades de cinta magnética u ópticas, así como capacidades de comunicaciones de alta velocidad y potentes CPUS, normalmente específicas para sus cometidos.

Estaciones de trabajo (*Workstation*) son computadoras de escritorio muy potentes destinadas a los usuarios pero con capacidades matemáticas y gráficas superiores a un PC y que pueden realizar tareas más complicadas que un PC en la misma o menor cantidad de tiempo. Tienen capacidad para ejecutar programas técnicos y cálculos científicos, y suelen utilizar UNIX o Windows NT como sistema operativo.

Las **minicomputadoras**, hoy día muchas veces confundidas con los servidores, son computadoras de rango medio, que se utilizan en centros de investigación, departamentos científicos, fábricas, etc., y que poseen una gran capacidad de proceso numérico y tratamiento de gráficos, fundamentalmente, aunque también son muy utilizadas en el mundo de la gestión, como es el caso de los conocidos AS/400 de IBM.

Las **grandes computadoras** (*mainframes*) son máquinas de gran potencia de proceso y extremadamente rápidas y además disponen de una gran capacidad de almacenamiento masivo. Son las grandes computadoras de los bancos, universidades, industrias, etc. Las **supercomputadoras**⁶ son las más potentes y sofisticadas que existen en la actualidad; se utilizan para tareas que requieren cálculos complejos y extremadamente rápidos. Estas computadoras utilizan numerosos procesadores en paralelo y tradicionalmente se han utilizado y utilizan para fines científicos y militares en aplicaciones tales como meteorología, previsión de desastres naturales, balística, industria aeroespacial, satélites, aviónica, biotecnología, nanotecnología, etc. Estas computadoras emplean numerosos procesadores en paralelo y se están comenzando a utilizar en negocios para manipulación masiva de datos. Una supercomputadora, ya popular es el *Blue Gene* de IBM o el *Mare Nostrum* de la Universidad Politécnica de Cataluña.

Además de esta clasificación de computadoras, existen actualmente otras microcomputadoras (*handheld computers*, computadoras de mano) que se incorporan en un gran número de dispositivos electrónicos y que constituyen el corazón y brazos de los mismos, por su gran capacidad de proceso. Este es el caso de los **PDA** (Asistentes Personales Digitales) que en muchos casos vienen con versiones específicas para estos dispositivos de los sistemas operativos populares, como es el caso de Windows Mobile, y en otros casos utilizan sistemas operativos exclusivos como es el caso de Symbian y Palm OS. También es cada vez más frecuente que otros dispositivos de mano, tales como los **teléfonos inteligentes**, **cámaras de fotos**, **cámaras digitales**, **videocámaras**, etc., incorporen tarjetas de memoria de 128 Mb hasta 4 GB, con tendencia a aumentar.

⁵ En muchos países de Latinoamérica, el término **computadora portátil**, es más conocido popularmente por su nombre en inglés, **laptop**.

⁶ En España existen varias supercomputadoras. A destacar, las existentes en el Centro de Supercomputación de Galicia, la de la Universidad Politécnica de Valencia y la de la Universidad Politécnica de Madrid. En agosto de 2004 se puso en funcionamiento en Barcelona, en la sede de la Universidad Politécnica de Cataluña, otra gran supercomputadora, en este caso de IBM que ha elegido España y, en particular Barcelona, como sede de esta gran supercomputadora que a la fecha de la inauguración se prevé esté entre las cinco más potentes del mundo. Esta supercomputadora denominada *Mare Nostrum* es una de las más potentes del mundo y está ubicada en el Centro de Supercomputación de Barcelona, dirigido por el profesor Mateo Valero, catedrático de Arquitectura de Computadoras de la Universidad Politécnica de Cataluña.

1.2. ORGANIZACIÓN FÍSICA DE UNA COMPUTADORA

Los dos componentes principales de una computadora son: *hardware* y *software*. **Hardware** es el equipo físico o los dispositivos asociados con una computadora. Sin embargo, para ser útil una computadora necesita además del equipo físico, un conjunto de instrucciones dadas. El conjunto de instrucciones que indican a la computadora aquello que deben hacer se denomina **software** o **programas** y se escriben por **programadores**. Este libro se centra en la enseñanza y aprendizaje de la programación o proceso de escribir programas.

Una **red** consta de un número de computadoras conectadas entre sí directamente o a través de otra computadora central (llamada *servidor*), de modo que puedan compartir recursos tales como impresoras, unidades de almacenamiento, etc., y que pueden compartir información. Una red puede contener un núcleo de PC, estaciones de trabajo y una o más computadoras grandes, así como dispositivos compartidos como impresora.

La mayoría de las computadoras, grandes o pequeñas, están organizadas como se muestra en la Figura 1.2. Una computadora consta fundamentalmente de cinco componentes principales: *dispositivos de entrada*; *dispositivos de salida*; *unidad central de proceso (UCP)* o *procesador* (compuesto de la **UAL**, Unidad Aritmética y Lógica y la **UC**, Unidad de Control); la *memoria principal o central*; *memoria secundaria o externa* y el *programa*.

Si a la organización física de la Figura 1.2 se le añaden los dispositivos para comunicación exterior con la computadora, aparece la estructura típica de un sistema de computadora que, generalmente, consta de los siguientes dispositivos de *hardware*:

- Unidad Central de Proceso, **UCP (CPU, Central Processing Unit)**.
- Memoria principal.
- Memoria secundaria (incluye medios de almacenamiento masivo como disquetes, memorias USB, discos duros, discos CD-ROM, DVD...).
- Dispositivos de entrada tales como teclado y ratón.
- Dispositivos de salida tales como monitores o impresoras.
- Conexiones de redes de comunicaciones, tales como módems, conexión *Ethernet*, conexiones **USB**, conexiones serie y paralelo, conexión *Firewire*, etc.

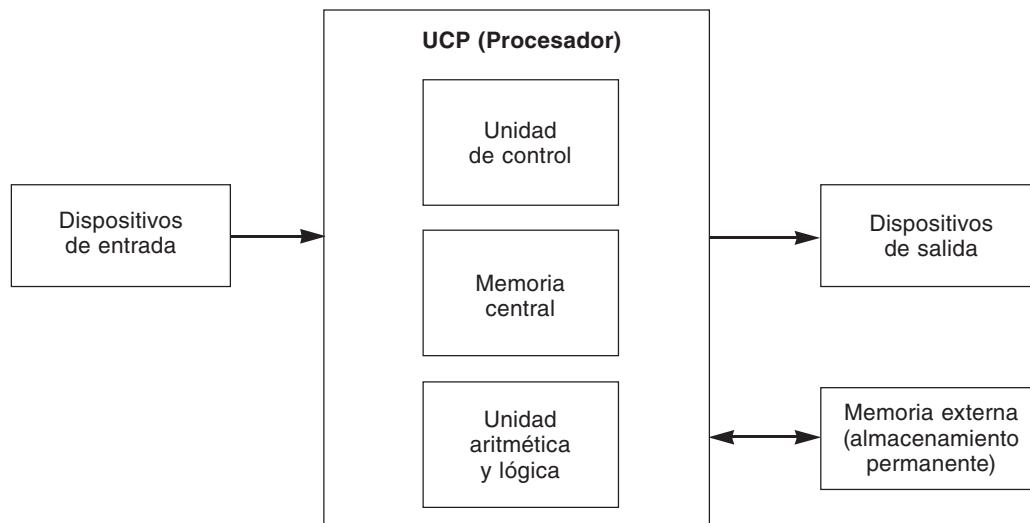


Figura 1.2. Organización física de una computadora.

Las computadoras sólo entienden un lenguaje compuesto únicamente por ceros y unos. Esta forma de comunicación se denomina **sistema binario** digital y en el caso concreto de las máquinas computadoras, código o **lenguaje máquina**. Este lenguaje máquina utiliza secuencias o patrones de ceros y unos para componer las instrucciones que posteriormente reciben de los diferentes dispositivos de la computadora, tales como el microprocesador, las unidades de discos duros, los teclados, etc.

La Figura 1.2 muestra la integración de los componentes que conforman una computadora cuando se ejecuta un programa; las flechas conectan los componentes y muestran la dirección del flujo de información.

El **programa** se debe transferir primero de la *memoria secundaria* a la *memoria principal* antes de que pueda ser ejecutado. Los datos se deben proporcionar por alguna fuente. La persona que utiliza un programa (**usuario** de programa) puede proporcionar datos a través de un dispositivo de entrada. Los datos pueden proceder de un **archivo (fichero)**, o pueden proceder de una máquina remota vía una conexión de red de la empresa o bien la red Internet.

Los datos se almacenan en la **memoria principal** de una computadora a la cual se puede acceder y manipular mediante la **unidad central de proceso (UCP)**. Los resultados de esta manipulación se almacenan de nuevo en la memoria principal. Por último, los resultados (la información) de la memoria principal se pueden visualizar en un **dispositivo de salida**, guardar en un almacenamiento secundario o enviarse a otra computadora conectada con ella en red.

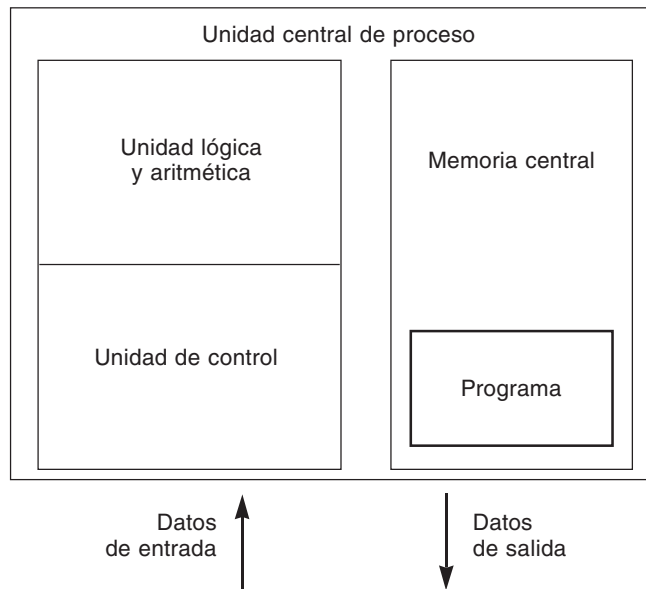


Figura 1.3. Unidad Central de Proceso.

Uno de los componentes fundamentales de un PC es la placa base (en inglés, *motherboard* o *mainboard*) que es una gran placa de circuito impreso que conecta entre sí los diferentes elementos contenidos en ella y sobre la que se conectan los elementos más importantes del PC: zócalo del microprocesador, zócalos de memoria, diferentes conectores, ranuras de expansión, puertos, etc.

Los paquetes de datos (de 8, 16, 32, 64 o más bits a la vez) se mueven continuamente entre la CPU y todos los demás componentes (memoria RAM, disco duro, etc.). Estas transferencias se realizan a través de *buses*. Los **buses** son los canales de datos que interconectan los componentes del PC; algunos están diseñados para transferencias pequeñas y otros para transferencias mayores. Existen diferentes *buses* siendo el más importante el *bus frontal (FSB, Front Side Bus)* en los sistemas actuales o *bus del sistema* (en sistemas más antiguos) y que conectan la CPU o procesador con la memoria RAM. Otros *buses* importantes son los que conectan la placa base de la computadora con los dispositivos periféricos del PC y se denominan *buses de E/S*.

1.2.1. Dispositivos de Entrada/Salida (E/S): periféricos

Los dispositivos de *Entrada/Salida (E/S)* [*Input/Output (I/O)* en inglés] permiten la comunicación entre la computadora y el usuario. *Los dispositivos de entrada*, como su nombre indica, sirven para introducir datos (información) en la computadora para su proceso. Los datos se *leen* de los dispositivos de entrada y se almacenan en la memoria central o interna. Los dispositivos de entrada convierten la información de entrada en señales eléctricas que se almacenan en la memoria central. Dispositivos de entrada típicos son los **teclados**; otros son: **lectores de tarjetas** —ya en desuso—, **lápices ópticos**, **palancas de mando (joystick)**, **lectores de códigos de barras**, **escáneres**, **micrófonos**, etc.

Hoy día tal vez el dispositivo de entrada más popular es el **ratón** (*mouse*) que mueve un puntero electrónico sobre la pantalla que facilita la interacción usuario-máquina⁷.

Los *dispositivos de salida* permiten representar los resultados (salida) del proceso de los datos. El dispositivo de salida típico es la **pantalla (CRT)**⁸ o **monitor**. Otros dispositivos de salida son: **impresoras** (imprimen resultados en papel), **trazadores gráficos** (*plotters*), **reconocedores de voz**, **altavoces**, etc.

El teclado y la pantalla constituyen —en muchas ocasiones— un único dispositivo, denominado **terminal**. Un teclado de terminal es similar al teclado de una máquina de escribir moderna con la diferencia de algunas teclas extras que tiene el terminal para funciones especiales. Si está utilizando una computadora personal, el teclado y el monitor son dispositivos independientes conectados a la computadora por cables. En ocasiones, la impresora se conoce como **dispositivo de copia dura** (*hard copy*), debido a que la escritura en la impresora es una copia permanente (dura) de la salida, y en contraste a la pantalla se la denomina **dispositivo de copia blanda** (*soft copy*), ya que la pantalla actual se pierde cuando se visualiza la siguiente.

Los dispositivos de entrada/salida y los dispositivos de almacenamiento secundario o auxiliar (memoria externa) se conocen también con el nombre de *dispositivos periféricos* o simplemente **periféricos** ya que, normalmente, son externos a la computadora. Estos dispositivos son unidades de discos [disquetes (ya en desuso), CD-ROM, DVD, cintas, etc.], videocámaras, teléfonos celulares (móviles), etc. Todos los dispositivos periféricos se conectan a las computadoras a través de conectores y **puertos** (*ports*) que son interfaces electrónicos.

1.2.2. La memoria principal

La memoria de una computadora almacena los datos de entrada, programas que se han de ejecutar y resultados. En la mayoría de las computadoras existen dos tipos de memoria principal: **memoria de acceso aleatorio RAM** que soporta almacenamiento temporal de programas y datos y **memoria de sólo lectura ROM** que almacena datos o programas de modo permanente.

La **memoria central (RAM, Random, Access Memory)** o simplemente **memoria** se utiliza para almacenar, de modo temporal información, datos y programas. En general, la información almacenada en memoria puede ser de dos tipos: las *instrucciones* de un programa y los *datos* con los que operan las instrucciones. Para que un programa se pueda *ejecutar* (correr, rodar, funcionar..., en inglés *run*), debe ser situado en la memoria central, en una operación denominada *carga* (*load*) del programa. Después, cuando se ejecuta (se realiza, funciona) el programa, *cualquier dato a procesar por el programa se debe llevar a la memoria* mediante las instrucciones del programa. En la memoria central, hay también datos diversos y espacio de almacenamiento temporal que necesita el programa cuando se ejecuta y así poder funcionar⁹.

La memoria principal es la encargada de almacenar los programas y datos que se están ejecutando y su principal característica es que el acceso a los datos o instrucciones desde esta memoria es muy rápido.

Es un tipo de memoria volátil (su contenido se pierde cuando se apaga la computadora); esta memoria es, en realidad, la que se suele conocer como memoria principal o de trabajo; en esta memoria se pueden escribir datos y leer de ella. Esta memoria RAM puede ser *estática (SRAM)* o *dinámica (DRAM)* según sea el proceso de fabricación. Las memorias RAM actuales más utilizadas son las **SDRAM** en sus dos tipos: **DDR** (Double Data Rate) y **DDR2**.

En la memoria principal se almacenan:

- Los datos enviados para procesarse desde los dispositivos de entrada.
- Los programas que realizarán los procesos.
- Los resultados obtenidos preparados para enviarse a un dispositivo de salida.

La memoria **ROM**, es una memoria que almacena información de modo permanente en la que no se puede escribir (viene pregrabada “grabada” por el fabricante) ya que es una *memoria de sólo lectura*. Los programas alma-

⁷ Todas las acciones a realizar por el usuario se realizarán con el ratón con la excepción de las que requieren de la escritura de datos por teclado. El nombre de ratón parece que proviene de la similitud del cable de conexión con la cola de un ratón. Hoy día, sin embargo, este razonamiento carece de sentido ya que existen ratones inalámbricos que no usan cable y se comunican entre sí a través de rayos infrarrojos.

⁸ *Cathode Ray Tube*: Tubo de rayos catódicos.

⁹ En la jerga informática también se conoce esta operación como “correr un programa”.

cenados en ROM no se pierden al apagar la computadora y cuando se enciende, se lee la información almacenada en esta memoria. Al ser esta memoria de sólo lectura, los programas almacenados en los chips ROM no se pueden modificar y suelen utilizarse para almacenar los programas básicos que sirven para arrancar la computadora.

Con el objetivo de que el procesador pueda obtener los datos de la memoria central más rápidamente, la mayoría de los procesadores actuales (muy rápidos) utilizan con frecuencia una *memoria* denominada **caché** que sirva para almacenamiento intermedio de datos entre el procesador y la memoria principal. La memoria caché —en la actualidad— se incorpora casi siempre al procesador.

Los programas y los datos se almacenan en RAM. Las memorias de una computadora personal se miden en unidades de memoria (se describen en el apartado 1.2.3) y suelen ser actualmente de 512 MB a 1, 2 o 3 GB, aunque ya es frecuente encontrar memorias centrales de 4 y 8 GB en computadoras personales y en cantidad mayor en computadoras profesionales y en servidores.

Normalmente una computadora contiene mucha más memoria RAM que memoria ROM interna; también la cantidad de memoria se puede aumentar hasta un máximo especificado, mientras que la cantidad de memoria ROM, normalmente es fija. Cuando en la jerga informática y en este texto se menciona la palabra memoria se suele referir a memoria RAM que normalmente es la memoria accesible al programador.

La memoria RAM es una memoria muy rápida y limitada en tamaño, sin embargo la computadora tiene otro tipo de memoria denominada *memoria secundaria o almacenamiento secundario* que puede crecer comparativamente en términos mucho mayores. La **memoria secundaria** es realmente un dispositivo de almacenamiento masivo de información y por ello, a veces, se la conoce como *memoria auxiliar, almacenamiento auxiliar, almacenamiento externo y memoria externa*.

1.2.3. Unidades de medida de memoria

La **memoria principal** es uno de los componentes más importantes de una computadora y sirve para almacenamiento de información (datos y programas). Existen dos tipos de memoria y de almacenamiento: Almacenamiento principal (memoria principal o memoria central) y almacenamiento secundario o almacenamiento masivo (discos, cintas, etc.).

La memoria central de una computadora es una zona de almacenamiento organizada en centenares o millares de unidades de almacenamiento individual o celdas. La memoria central consta de un conjunto de *celdas de memoria* (estas **celdas o posiciones de memoria** se denominan también *palabras*, aunque no “guardan” analogía con las palabras del lenguaje). Cada palabra puede ser un grupo de 8 bits, 16 bits, 32 bits o incluso 64 bits, en las computadoras más modernas y potentes. Si la palabra es de 8 bits se conoce como *byte*. El término *bit (dígito binario)*¹⁰ se deriva de las palabras inglesas “**binary digit**” y es la unidad de información más pequeña que puede tratar una computadora. El término *byte* es muy utilizado en la jerga informática y, normalmente, las palabras de 16 bits se suelen conocer como palabras de 2 *bytes*, y las palabras de 32 bits como palabras de 4 *bytes*.

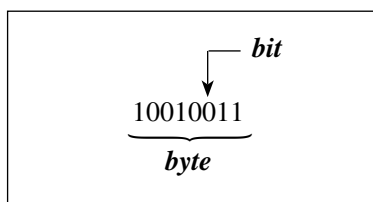


Figura 1.4. Relación entre un bit y un byte.

La memoria central de una computadora puede tener desde unos centenares de millares de *bytes* hasta millones de *bytes*. Como el *byte* es una unidad elemental de almacenamiento, se utilizan múltiplos para definir el tamaño de la memoria central: **Kilobyte (KB)** igual a 1.024 *bytes*¹¹ (2^{10}), **Megabyte (MB)** igual a 1.024×1.024 *bytes* ($2^{20} = 1.048.576$),

¹⁰ Binario se refiere a un sistema de numeración basado en los dos números o dígitos, 0 y 1; por consiguiente, un bit es o bien un 0 o bien un 1.

¹¹ Se adoptó el término **Kilo** en computadoras debido a que 1.024 es muy próximo a 1.000, y por eso en términos familiares y para que los cálculos se puedan hacer fáciles mentalmente se asocia 1 KB a 1.000 *bytes* y 1 MB a 1.000.000 de *bytes* y 1 GB a 1.000.000.000 de *bytes*. Así, cuando se habla en jerga diaria de 5 KB estamos hablando, en rigor, de $5 \times 1.024 = 5.120$ *bytes*, pero en cálculos consideramos 5.000 *bytes*. De

Gigabyte (GB) igual a 1.024 MB ($2^{30} = 1.073.741.824$). Las abreviaturas **MB** y **GB** se han vuelto muy populares como unidades de medida de la potencia de una computadora.

Desgraciadamente la aplicación de estos prefijos representa un mal uso de la terminología de medidas, ya que en otros campos las referencias a las unidades son potencias de 10. Por ejemplo, las medidas en distancias, *Kilómetro (Km)* se refiere a 1.000 metros, las medidas de frecuencias, *Megahercio (MHz)* se refieren a 1.000.000 de hercios. En la jerga informática popular para igualar terminología, se suele hablar de 1 **KB** como 1.000 bytes y 1 **MB** como 1.000.000 de bytes y un 1 **GB** como 1.000 millones de bytes, sobre todo para correspondencia y fáciles cálculos mentales, aunque como se observa en la Tabla 1.1 estos valores son sólo aproximaciones prácticas.

Tabla 1.1. Unidades de medida de almacenamiento

Byte	Byte (B)	<i>equivale a</i>	8 bits	
Kilobyte	Kbyte (KB)	<i>equivale a</i>	1.024 bytes	(10^3)
Megabyte	Mbyte (MB)	<i>equivale a</i>	1.024 Kbytes	(10^6)
Gigabyte	Gbyte (GB)	<i>equivale a</i>	1.024 Mbytes	(10^9)
Terabyte	Tbyte (TB)	<i>equivale a</i>	1.024 Gbytes	(10^{12})
Petabyte	Pbyte (PB)	<i>equivale a</i>	1.024 Tbytes	(10^{15})
Exabyte	Ebyte (EB)	<i>equivale a</i>	1.024 Pbytes	(10^{18})
Zettabyte	Zbyte (ZB)	<i>equivale a</i>	1.024 Ebytes	(10^{21})
Yotta	Ybyte (YB)	<i>equivale a</i>	1.024 Zbytes	(10^{24})

1 Tb = 1.024 Gb; 1 GB = 1.024 Mb = 1.048.576 Kb = 1.073.741.824 b

Celda de memoria

- La memoria de una computadora es una secuencia ordenada de celdas de memoria.
- Cada celda de memoria tiene una única dirección que indica su posición relativa en la memoria.
- Los datos se almacenan en una celda de memoria y constituyen el contenido de dicha celda.

Byte

Un **byte** es una posición de memoria que puede contener ocho *bits*. Cada bit sólo puede contener dos valores posibles, 0 o 1. Se requieren ocho bits (un *byte*) para codificar un carácter (una letra u otro símbolo del teclado).

Bytes, direcciones, memoria

La memoria principal se divide en posiciones numeradas que se denominan **bytes**. A cada *byte* se asocia un número denominado **dirección**. Un número o una letra se representan por un grupo de *bytes* consecutivos en una posición determinada. La dirección del primer *byte* del grupo se utiliza como la dirección más grande de esta posición de memoria.

Espacio de direccionamiento

Para tener acceso a una palabra en la memoria se necesita un identificador que a nivel de *hardware* se le conoce como dirección. Existen dos conceptos importantes asociados a cada celda o posición de memoria: su **dirección** y su **contenido**. Cada celda o *byte* tiene asociada una única *dirección* que indica su posición relativa en memoria y mediante

este modo se guarda correspondencia con las restantes representaciones de las palabras Kilo, Mega, Giga... Usted debe considerar siempre los valores reales para 1 KB, 1 MB o 1 GB, mientras esté en su fase de formación y posteriormente en el campo profesional desde el punto de vista de programación, para evitar errores técnicos en el diseño de sus programas, y sólo recurrir a las cifras mil, millón, etc., para la jerga diaria.

la cual se puede acceder a la posición para almacenar o recuperar información. La información almacenada en una posición de memoria es su contenido. La Figura 1.5 muestra una memoria de computadora que consta de 1.000 posiciones en memoria con direcciones de 0 a 999 en código decimal. El contenido de estas direcciones o posiciones de memoria se llaman **palabras**, que como ya se ha comentado pueden ser de 8, 16, 32 y 64 bits. Por consiguiente, si trabaja con una máquina de 32 bits, significa que en cada posición de memoria de su computadora puede alojar 32 bits, es decir 32 dígitos, bien ceros o unos.

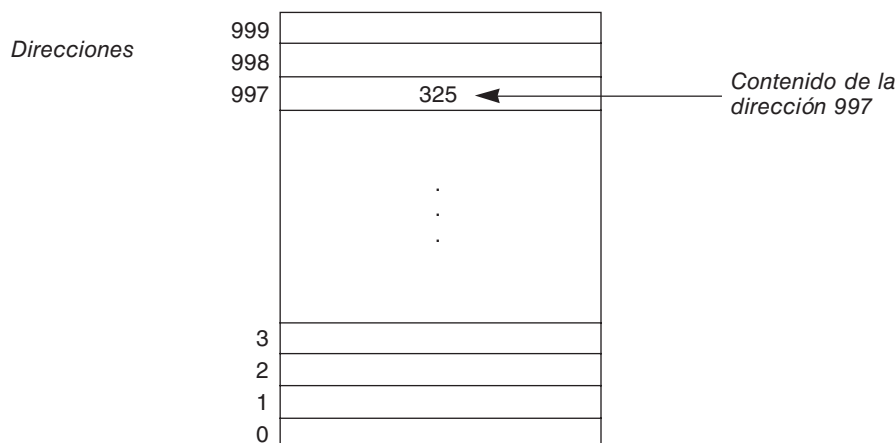


Figura 1.5. Memoria central de una computadora.

Las direcciones de memoria se definen usando enteros binarios sin signo o sus correspondientes enteros decimales.

El número de posiciones únicas identificables en memoria se denomina **espacio de direccionamiento**. Por ejemplo, en una memoria de 64 *kilobytes* (KB) y un tamaño de palabra de un *byte* tienen un espacio de direccionamiento que varía de 0 a 65.535 (64 KB, $64 \times 1.024 = 65.536$).

Los bytes sirven para representar los caracteres (letras, números y signos de puntuación adicionales) en un código estándar internacional denominado **ASCII** (**A**merican **S**tandard **C**ode for **I**nformation **I**nterchange), utilizado por todas las computadoras del mundo, o bien en un código estándar más moderno denominado **Unicode**. Todos estos símbolos se almacenan en memoria y con ellos trabajan las computadoras.

1.2.4. El procesador

El **procesador** o **Unidad Central de Proceso, UCP (CPU, Central Processing Unit)** controla el funcionamiento de la computadora y realiza sus funciones de procesamiento de los datos, constituyendo el cerebro y corazón de la computadora o también su sistema nervioso. Se encarga de un modo práctico de realizar numerosos cálculos y operaciones ordenadas por los diferentes programas instalados en la computadora.

Cada computadora tiene al menos una UCP para interpretar y ejecutar las instrucciones de cada programa, realizar las manipulaciones de datos aritméticos y lógicos, y comunicarse con todas las restantes partes de la máquina indirectamente a través de la memoria.

Un moderno procesador o microprocesador, es una colección compleja de dispositivos electrónicos. En una computadora de escritorio o en una portátil (laptop o *notebook*) la UCP se aloja junto con otros *chips* y componentes electrónicos en la placa base también denominada placa madre (*motherboard*). La elección de la placa base proporcionará una mayor o menor potencia a la computadora y está compuesta por numerosos componentes electrónicos y se ramifica hacia todos los periféricos externos a través de conectores (puertos) colocados en la mayoría de las veces en la parte posterior del equipo, principalmente en los equipos de sobremesa y torre, mientras que en los equipos portátiles o portables, están colocados no sólo en la parte posterior sino también en las partes laterales o incluso delantera.

Existen numerosos fabricantes de procesadores aunque, entre otros, los más acreditados son Intel, AMD, Transmeta (empresa conocida por estar vinculada en sus orígenes con Linus Torvald creador del sistema operativo Linux), IBM, Motorola y Sun Microsystems. En cuanto a familias en concreto, los más populares son: Pentium de Intel (que

incluye Celeron y Xeon), Opteron de AMD, SPARC de Sun Microsystems, Crusoe de Transmeta, Centrino Core 2 y Centro Core 2 Duo de Intel que se instalan en portátiles, etc.

Todas las UCP tienen una velocidad de trabajo, regulada por un pequeño *crystal de cuarzo*, y que se conoce como *frecuencia de reloj*. El cristal vibra a un elevado número de ciclos de reloj. Con cada ciclo de reloj se envía un impulso a la UCP, y en principio, cada pulsación puede hacer realizar una o más tareas a la UCP. El número de ciclos de reloj por segundo se mide en hertzios. El cristal de la UCP vibra millones de veces por segundo y por esta razón la velocidad del reloj se calcula en millones de oscilaciones (megahercios o MHz) o miles de millones de ciclos por segundo, gigahercios (GHz). En consecuencia la velocidad de los microprocesadores se mide en MHz o en GHz. De esta forma si el procesador de su equipo funciona a 3 GHz significa que realiza 3 millones de operaciones por segundo.

Generaciones de microprocesadores

El PC original de 1981 trabajaba a 4,77 MHz y su microprocesador era el Intel 8088. Trabajaba a 16 bits internamente, aunque el bus externo para comunicarse con el resto de componentes era tan sólo de 8 bits. El microprocesador Intel 8088 fue lanzado al mercado en junio de 1979, aunque con anterioridad (junio de 1978) Intel lanzó el 8086. Estos microprocesadores con sus diferentes modelos, constituyeron la primera generación o familia de microprocesadores. En total, Intel ha lanzado numerosas generaciones o familias de procesadores que han permanecido en el mercado durante varios años durante los cuales se ha ido incrementando la frecuencia de reloj.

En 1993 Intel presentó el Pentium II, Motorola el 68060 y AMD el K5. Desde entonces Intel y AMD, fundamentalmente, han continuado presentando numerosas generaciones o familias de procesadores que permanecen en el mercado durante varios años incrementando la frecuencia de reloj con cada nuevo modelo además de otras características importantes. En el año 2000 Intel presentó el Pentium IV y AMD el Athlon XP y Duron, desencadenantes de los potentes procesadores existentes hoy día y que han servido de soporte a la mayoría de las computadoras personales de la primera década de los 2000. En 2004, Intel presentó los **Pentium M, D** y **Core Duo**, mientras que AMD presentó en 2005, el **AMD Athlon**.

En enero de 2006, Intel lanzó el procesador **Core Duo**, optimizado para aplicaciones de procesos múltiples y multitarea. Puede ejecutar varias aplicaciones complejas simultáneamente, como juegos con gráficos potentes o programas que requieran muchos cálculos y al mismo tiempo puede descargar música o analizar su PC con un antivirus en segundo plano. A finales del mismo año, Intel presentó el **Core 2 Duo** que dobla la potencia de cálculo y reduce considerablemente el consumo de energía. Intel sigue fabricando para sus equipos de sobremesa y portátiles, procesadores Pentium (Pentium D y Pentium 4) y procesadores Celeron

En 2007 han aparecido los procesadores de más de un núcleo, tales como **Intel Core 2 Quad**, **AMD Quad Core** y **AMD Quad FX**, todos ellos de cuatro núcleos y en 2008 se espera el lanzamiento de procesadores Intel y AMD con más de ocho núcleos. De igual modo Intel también fabrica Pentium de dos y cuatro núcleos. (Pentium Dual Core, Pentium Cuad Core) y Athlon 64 y con tendencia a aumentar el número de núcleos.

Proceso de ejecución de un programa

El ratón y el teclado introducen datos en la memoria central cuando se ejecuta el programa. Los datos intermedios o auxiliares se transfieren desde la unidad de disco a la pantalla y a la unidad de disco, a medida que se ejecuta el programa.

Cuando un programa se ejecuta, se debe situar primero en memoria central de igual modo que los datos. Sin embargo, la información almacenada en la memoria se pierde (borra) cuando se *apaga* (desconecta de la red eléctrica) la computadora, y por otra parte la memoria central es limitada en capacidad. Por esta razón, para poder disponer de almacenamiento permanente, tanto para programas como para datos, se necesitan *dispositivos de almacenamiento secundario, auxiliar o masivo (mass storage, o secondary storage)*.

En el campo de las computadoras es frecuente utilizar la palabra **memoria** y almacenamiento o **memoria** externa, indistintamente. En este libro —y recomendamos su uso— se utilizará el término memoria sólo para referirse a la memoria central.

Comparación de la memoria central y la memoria auxiliar

La memoria central o principal es mucho más rápida y cara que la memoria auxiliar. Se deben transferir los datos desde la memoria auxiliar hasta la memoria central, antes de que puedan ser procesados. Los datos en memoria central son: volátiles y desaparecen cuando se apaga la computadora. Los datos en memoria auxiliar son permanentes y no desaparecen cuando se apaga la computadora.

1.2.5. Propuestas para selección de la computadora ideal para aprender programación o para actividades profesionales

Las computadoras personales que en el primer trimestre de 2008 se comercializan para uso doméstico y en oficinas o en las empresas suelen tener características comunes, y es normal que sus prestaciones sean similares a las utilizadas en los laboratorios de programación de Universidades, Institutos Tecnológicos y Centros de Formación Profesional. Por estas razones en la Tabla 1.2 se incluyen recomendaciones de características técnicas medias que son, normalmente, utilizadas, para prácticas de aprendizaje de programación, por el alumno o por el lector autodidacta, así como por profesionales en su actividad diaria.

Tabla 1.2. Características técnicas recomendadas para computadoras de escritorio (profesionales y uso doméstico)

Procesador (Computadora de sobremesa-computadora portátil o <i>laptop</i>)	Intel	
	www.intel.com/cd/products/services/emea/spa/processors/322163.htm*	
	Procesadores para equipos de sobremesa	
	<ul style="list-style-type: none"> • Intel Core 2 Extreme • Intel Core 2 Quad • Intel Core 2 Duo • Intel Celeron, Celeron D y Celeron de doble núcleo 	<ul style="list-style-type: none"> • Pentium Extreme • Pentium D Edition • Pentium 4
	Procesadores para portátiles (<i>laptop</i>)	
<ul style="list-style-type: none"> • Intel Centrino Duo con procesador Intel Core Duo • Intel Pentium e Intel Celeron 		
	AMD	
	www.amd.com/es-es/Processors/ProductInformation/0,,30_118,00.html*	
	Procesadores para equipos de sobremesa	
	<ul style="list-style-type: none"> • AMD Phenom Quad Core • AMD Athlon 	<ul style="list-style-type: none"> • AMD Sempron
Procesadores para portátiles (<i>laptop</i>)		
	<ul style="list-style-type: none"> • Tecnología Mobile AMD Turion 64 × 2 Dual Core 	<ul style="list-style-type: none"> • Mobile Athlon 64 × 2 • Mobile AMD Sempron
Memoria RAM	512 MB, 1 GB a 4 GB DDR2	
Disco duro	SATA 80 GB, 160 G, 250 GB, 320 GB y superiores	
Tarjeta gráfica	Memoria dedicada o compartida, 256-1.024 Mb Nvidia GeForce ATI Mobility Radeom Intel	
Grabadora	DVD +/- RW de doble capa <i>Blue-Ray</i> HD-DVD (desde febrero de 2008, se ha dejado de comercializar por Toshiba)	
Pantalla	7", 11,1", 11,9", 12,1", 13", 13,3", 14", 15", 15,4", 17", 19", 20"	
Sistema operativo	Windows XP, Mac OS Windows Vista Home, Premiun, Business, Ultimate Mac OS Linux	
Redes y conectividad	<i>Wifi</i> <i>Bluetooth</i> v2.0 LAN USB 2.0	Protocolos <i>a, b, g, n</i> Ethernet 10/100, 10/100/1000 Varios puertos (2 o más)
Otras características	<i>WebCam</i> de 1,3-2 Mpixel Sintonizadora TV GPS integrado Lector multitarjetas <i>Firewire</i> 3G (UMTS), 3,5 G (HSDPA), 3,75 G (HSUPA), HSPA (ya existen en el mercado computadoras con banda ancha móvil —HSPA— integrada, p.e. modelo DELL Vostro 1500)	

* En estas direcciones Web, el lector encontrará todas las especificaciones y características técnicas más sobresalientes de los procesadores de los fabricantes Intel y AMD.

1.3. REPRESENTACIÓN DE LA INFORMACIÓN EN LAS COMPUTADORAS

Una computadora es un sistema para procesar información de modo automático. Un tema vital en el proceso de funcionamiento de una computadora es estudiar la forma de representación de la información en dicha computadora. Es necesario considerar cómo se puede codificar la información en patrones de bits que sean fácilmente almacenables y procesables por los elementos internos de la computadora.

Las formas de información más significativas son: textos, sonidos, imágenes y valores numéricos y, cada una de ellas presentan peculiaridades distintas. Otros temas importantes en el campo de la programación se refieren a los métodos de detección de errores que se puedan producir en la transmisión o almacenamiento de la información y a las técnicas y mecanismos de comprensión de información al objeto de que ésta ocupe el menor espacio en los dispositivos de almacenamiento y sea más rápida su transmisión.

1.3.1. Representación de textos

La información en formato de texto se representa mediante un código en el que cada uno de los distintos símbolos del texto (tales como letras del alfabeto o signos de puntuación) se asignan a un único patrón de bits. El texto se representa como una cadena larga de bits en la cual los sucesivos patrones representan los sucesivos símbolos del texto original.

En resumen, se puede representar cualquier información escrita (texto) mediante caracteres. Los caracteres que se utilizan en computación suelen agruparse en cinco categorías:

1. **Caracteres alfabéticos** (letras mayúsculas y minúsculas, en una primera versión del abecedario inglés).

A, B, C, D, E, ... X, Y, Z, a, b, c, ... , X, Y, Z

2. **Caracteres numéricos** (dígitos del sistema de numeración).

0, 1, 2, 3, 4, 5, 6, 7, 8, 9 *sistema decimal*

3. **Caracteres especiales** (símbolos ortográficos y matemáticos no incluidos en los grupos anteriores).

{ } Ñ ñ ! ? & > # ç ...

4. **Caracteres geométricos y gráficos** (símbolos o módulos con los cuales se pueden representar cuadros, figuras geométricas, iconos, etc.

| □ ▭ ♠ ~ ...

5. **Caracteres de control** (representan órdenes de control como el carácter para pasar a la siguiente línea [NL] o para ir al comienzo de una línea [RC, *retorno de carro*, “*carriage return*, **CR**”] emitir un pitido en el terminal [BEL], etc.).

Al introducir un texto en una computadora, a través de un periférico, los caracteres se codifican según un **código de entrada/salida** de modo que a cada carácter se le asocia una determinada combinación de n bits.

Los códigos más utilizados en la actualidad son: **EBCDIC**, **ASCII** y **Unicode**.

- **Código EBCDIC** (*Extended Binary Coded Decimal Inter Change Code*).

Este código utiliza $n = 8$ bits de forma que se puede codificar hasta $m = 2^8 = 256$ símbolos diferentes. Éste fue el primer código utilizado para computadoras, aceptado en principio por IBM.

- **Código ASCII** (*American Standard Code for Information Interchange*).

El código ASCII básico utiliza 7 bits y permite representar 128 caracteres (letras mayúsculas y minúsculas del alfabeto inglés, símbolos de puntuación, dígitos 0 a 9 y ciertos controles de información tales como retorno de carro, salto de línea, tabulaciones, etc.). Este código es el más utilizado en computadoras, aunque el ASCII ampliado con 8 bits permite llegar a 2^8 (256) caracteres distintos, entre ellos ya símbolos y caracteres especiales de otros idiomas como el español.

- **Código Unicode**

Aunque ASCII ha sido y es dominante en la representación de los caracteres, hoy día se requiere de la necesidad de representación de la información en muchas otras lenguas, como el portugués, español, chino, el japonés, el árabe, etc. Este código utiliza un patrón único de 16 bits para representar cada símbolo, que permite 2^{16} bits o sea hasta 65.536 patrones de bits (símbolos) diferentes.

Desde el punto de vista de unidad de almacenamiento de caracteres, se utiliza el archivo (**fichero**). Un **archivo** consta de una secuencia de símbolos de una determinada longitud codificados utilizando ASCII o Unicode y que se denomina **archivo de texto**. Es importante diferenciar entre archivos de texto simples que son manipulados por los programas de utilidad denominados **editores de texto** y los archivos de texto más elaborados que se producen por los procesadores de texto, tipo Microsoft Word. Ambos constan de caracteres de texto, pero mientras el obtenido con el editor de texto, es un archivo de texto puro que codifica carácter a carácter, el archivo de texto producido por un procesador de textos contiene números, códigos que representan cambios de formato, de tipos de fuentes de letra y otros, e incluso pueden utilizar códigos propietarios distintos de ASCII o Unicode.

1.3.2. Representación de valores numéricos

El almacenamiento de información como caracteres codificados es ineficiente cuando la información se registra como numérica pura. Veamos esta situación con la codificación del número 65; si se almacena como caracteres ASCII utilizando un byte por símbolo, se necesita un total de 16 bits, de modo que el número mayor que se podía almacenar en 16 bits (dos bytes) sería 99. Sin embargo, si utilizamos *notación binaria* para almacenar enteros, el rango puede ir de 0 a 65.535 ($2^{16} - 1$) para números de 16 bits. Por consiguiente, la notación binaria (o variantes de ellas) es la más utilizada para el almacenamiento de datos numéricos codificados.

La solución que se adopta para la representación de datos numéricos es la siguiente: al introducir un número en la computadora se codifica y se almacena como un texto o cadena de caracteres, pero dentro del programa a cada dato se le envía un tipo de dato específico y es tarea del programador asociar cada dato al tipo adecuado correspondiente a las tareas y operaciones que se vayan a realizar con dicho dato.

El método práctico realizado por la computadora es que una vez definidos los datos numéricos de un programa, una rutina (función interna) de la biblioteca del compilador (traductor) del lenguaje de programación se encarga de transformar la cadena de caracteres que representa el número en su notación binaria.

Existen dos formas de representar los datos numéricos: números enteros o números reales.

Representación de enteros

Los datos de tipo entero se representan en el interior de la computadora en notación binaria. La memoria ocupada por los tipos enteros depende del sistema, pero normalmente son dos, bytes (en las versiones de MS-DOS y versiones antiguas de Windows y cuatro bytes en los sistemas de 32 bits como Windows o Linux). Por ejemplo, un entero almacenado en 2 bytes (16 bits):

```
1000 1110 0101 1011
```

Los enteros se pueden representar con signo (*signed*, en C++) o sin signo (*unsigned*, en C++); es decir, números positivos o negativos. Normalmente, se utiliza un bit para el signo. Los enteros sin signo al no tener signo pueden contener valores positivos más grandes. Normalmente, si un entero no se especifica “con/sin signo” se suele asignar con signo por defecto u omisión.

El rango de posibles valores de enteros depende del tamaño en bytes ocupado por los números y si se representan con signo o sin signo (la Tabla 1.3 resume características de tipos estándar en C++).

Representación de reales

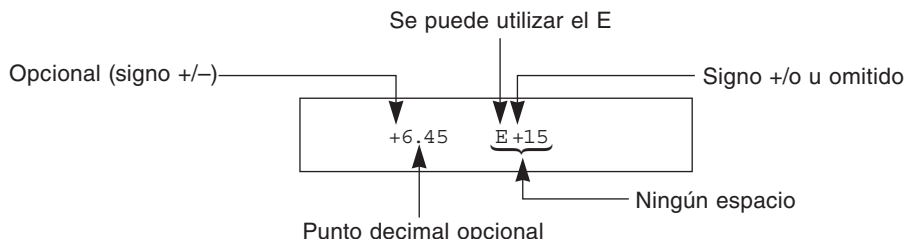
Los números reales son aquellos que contienen una parte decimal como 2,6 y 3,14152. Los reales se representan en *notación científica* o en *coma flotante*; por esta razón en los lenguajes de programación, como C++, se conocen como números en coma flotante.

Existen dos formas de representar los números reales. La primera se utiliza con la notación del punto decimal (*ojo* en el formato de representación español de números decimales, la parte decimal se representa por coma).

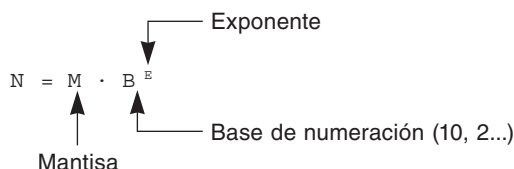
EJEMPLOS

12.35 99901.32 0.00025 9.0

La segunda forma para representar números en coma flotante en la notación científica o exponencial, conocida también como notación *E*. Esta notación es muy útil para representar números muy grandes o muy pequeños.



Notación exponencial



EJEMPLOS

2.52 e + 8	<i>equivale a</i>	252000000
8.34 E - 4	<i>equivale a</i>	8.34/10 ⁴ = 0.000834
7E5	<i>equivale a</i>	7000000
-18.35e15	<i>equivale a</i>	-18500000000000000
5.95E25	<i>equivale a</i>	595000000000000000000000000000
9.11e ₃₁	<i>equivale a</i>	0.000000000000000000000000000000911

Representación de caracteres

Un documento de texto se escribe utilizando un conjunto de caracteres adecuado al tipo de documento. En los lenguajes de programación se utilizan, principalmente, dos códigos de caracteres. El más común es **ASCII** (American Standard Code for Information Interchange) y algunos lenguajes, tal como Java, utilizan **Unicode** (www.unicode.org). Ambos códigos se basan en la asignación de un código numérico a cada uno de los tipos de caracteres del código.

En C++, los *caracteres* se procesan normalmente usando el tipo `char`, que asocia cada carácter a un código numérico que se almacena en un byte.

El código ASCII básico que utiliza 7 bits (128 caracteres distintos) y el ASCII *ampliado* a 8 bits (256 caracteres distintos) son los códigos más utilizados. Así se pueden representar caracteres tales como 'A', 'B', 'c', '\$', '4', '5', etc.

La Tabla 1.3, recoge los tipos enteros, reales y carácter utilizados en C++, la memoria utilizada (número de bytes ocupados por el dato) y el rango de números.

1.3.3. Representación de imágenes

Las imágenes se adquieren mediante periféricos especializados tales como escáneres, cámaras digitales de vídeo, cámaras fotográficas, etc. Una imagen, al igual que otros tipos de información, se representa por patrones de bits, generados por el periférico correspondiente. Existen dos métodos básicos para representar imágenes: *mapas de bits* y *mapas de vectores*.

Tabla 1.3. Tipos enteros reales, en C++

		Carácter y bool	
Tipo	Tamaño	Rango	
short (short int)	2 bytes	-32.738..32.767	
int	4 bytes	-2.147.483.648 a 2.147.483.647	
long (long int)	4 bytes	-2.147.483.648 a 2.147.483.647	
float (real)	4 bytes	10^{-38} a 10^{38} (aproximadamente)	
double	8 bytes	10^{-308} a 10^{308} (aproximadamente)	
long double	10 bytes	10^{-4932} a 10^{4932} (aproximadamente)	
char (carácter)	1 byte	Todos los caracteres ASCII	
bool	1 byte	True (verdadero) y false (falso)	

En las técnicas de *mapas de bits*, una imagen se considera como una colección de puntos, cada uno de los cuales se llama *pixel* (abreviatura de “*picture element*”). Una imagen en blanco y negro se representa como una cadena larga de bits que representan las filas de píxeles en la imagen, donde cada bit es bien 1 o bien 0, dependiendo de que el pixel correspondiente sea blanco o negro. En el caso de imágenes en color, cada pixel se representa por una combinación de bits que indican el color de los pixel. Cuando se utilizan técnicas de mapas de bits, el patrón de bits resultante se llama *mapa de bits*, significando que el patrón de bits resultante que representa la imagen es poco más que un mapa de la imagen.

Muchos de los periféricos de computadora —tales como cámaras de vídeo, escáneres, etc.— convierten imágenes de color en formato de mapa de bits. Los formatos más utilizados en la representación de imágenes se muestran en la Tabla 1.4.

Tabla 1.4. Mapas de bits

Formato	Origen y descripción
BMP	Microsoft. Formato sencillo con imágenes de gran calidad pero con el inconveniente de ocupar mucho (no útil para la web).
JPEG	Grupo JPEG. Calidad aceptable para imágenes naturales. Incluye compresión. Se utiliza en la web.
GIF	CompuServe. Muy adecuado para imágenes no naturales (logotipos, banderas, dibujos anidados...). Muy usado en la web.

Mapas de vectores. Otros métodos de representar una imagen se fundamentan en descomponer la imagen en una colección de objetos tales como líneas, polígonos y textos con sus respectivos atributos o detalles (grosor, color, etc.).

Tabla 1.5. Mapas de vectores

Formato	Descripción
IGES	ASME/ANSI. Estándar para intercambio de datos y modelos de (AutoCAD...).
Pict	Apple Computer. Imágenes vectoriales.
EPS	Adobe Computer.
TrueType	Apple y Microsoft para EPS.

1.3.4. Representación de sonidos

La representación de sonidos ha adquirido una importancia notable debido esencialmente a la infinidad de aplicaciones multimedia tanto autónomas como en la *web*.

El método más genérico de codificación de la información de audio para almacenamiento y manipulación en computadora es mostrar la amplitud de la onda de sonido en intervalos regulares y registrar las series de valores ob-

tenidos. La señal de sonido se capta mediante micrófonos o dispositivos similares y produce una señal analógica que puede tomar cualquier valor dentro de un intervalo continuo determinado. En un intervalo de tiempo continuo se dispone de infinitos valores de la señal analógica, que es necesario almacenar y procesar, para lo cual se recurre a una *técnica de muestreo*. Las muestras obtenidas se digitalizan con un conversor analógico-digital, de modo que la señal de sonido se representa por secuencias de bits (por ejemplo, 8 o 16) para cada muestra. Esta técnica es similar a la utilizada, históricamente, por las comunicaciones telefónicas a larga distancia. Naturalmente, dependiendo de la calidad de sonido que se requiera, se necesitarán más números de bits por muestra, frecuencias de muestreo más altas y lógicamente más muestreos por períodos de tiempo¹².

Como datos de referencia puede considerarse que para obtener reproducción de calidad de sonido de alta fidelidad para un disco CD de música, se suele utilizar, al menos, una frecuencia de muestreo de 44.000 muestras por segundo. Los datos obtenidos en cada muestra se codifican en 16 bits (32 bits para grabaciones en estéreo). Como dato anecdótico, cada segundo de música grabada en estéreo requiere más de un millón de bits.

Un sistema de codificación de música muy extendido en sintetizadores musicales es MIDI (*Musical Instruments Digital Interface*) que se encuentra en sintetizadores de música para sonidos de videojuegos, sitios web, teclados electrónicos, etc.

1.4. CODIFICACIÓN DE LA INFORMACIÓN

La información que manejan las computadoras es *digital*. Esto significa que esta información se construye a partir de unidades contables llamadas **dígitos**. Desde el punto de vista físico, las unidades de una computadora están constituidas por circuitos formados por componentes electrónicos denominados puertas, que manejan señales eléctricas que no varían de modo continuo sino que sólo pueden tomar dos estados discretos (dos voltajes). Cerrado y abierto, bajo y alto, 0 y 1. De este modo la memoria de una computadora está formada por millones de componentes de naturaleza digital que almacenan uno de dos estados posibles.

Una computadora no entiende palabras, números, dibujos ni notas musicales, ni incluso letras del alfabeto. De hecho, sólo entienden información que ha sido descompuesta en bits. Un *bit*, o dígito binario, es la unidad más pequeña de información que una computadora puede procesar. Un bit puede tomar uno de dos valores: 0 y 1. Por esta razón las instrucciones de la máquina y los datos se representan en códigos binarios al contrario de lo que sucede en la vida cotidiana en donde se utiliza el código o sistema decimal.

1.4.1. Sistemas de numeración

El sistema de numeración más utilizado en el mundo es el sistema decimal que tiene un conjunto de diez dígitos (0 al 9) y con la base de numeración 10. Así, cualquier número decimal se representa como una expresión aritmética de potencias de base 10; por ejemplo, 1.492, en base 10, se representa por la cantidad:

$$1492 = 1 \cdot 10^3 + 4 \cdot 10^2 + 9 \cdot 10^1 + 2 \cdot 10^0 = 1 \cdot 1000 + 4 \cdot 100 + 9 \cdot 10 + 2 \cdot 1$$

y 2.451,4 se representa por

$$2451,4 = 2 \cdot 10^3 + 4 \cdot 10^2 + 5 \cdot 10^1 + 1 \cdot 10^0 + 4 \cdot 10^{-1} = 2 \cdot 1000 + 4 \cdot 100 + 5 \cdot 10 + 1 \cdot 1 + 4 \cdot 0,1$$

Además del sistema decimal existen otros sistemas de numeración utilizados con frecuencia en electrónica e informática (computación): el sistema hexadecimal y el sistema octal.

El sistema o código **hexadecimal** tiene como base 16, y 16 dígitos para su representación (0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E y F), los diez dígitos decimales y las primeras letras del alfabeto que representan los dígitos de mayor peso, de valor 10, 11, 12, 13, 14 y 15.

El sistema o código **octal** tiene por base 8 y 8 dígitos (0, 1, 2, 3, 4, 5, 6 y 7).

En las computadoras, como ya se ha comentado, se utiliza el sistema binario o de base 2 con dos dígitos: 0 y 1. En el sistema de numeración binario o digital, cada número se representa por un único patrón de dígitos 0 y 1. La Tabla 1.6 representa los equivalentes de números en código decimal y binario.

¹² En las obras del profesor Alberto Prieto, Schaum “*Conceptos de Informática e Introducción a la Informática*”, publicadas en McGraw-Hill, puede encontrar una excelente referencia sobre estos conceptos y otros complementarios de este capítulo introductorio.

Tabla 1.6. Representación de números decimales y binarios

Representación decimal	Representación binaria
0	0
1	1
2	10
3	11
4	100
5	101
6	110
7	111
8	1000
9	1001
10	1010
11	1011
12	1100
13	1101
14	1110
15	1111

Así, un número cualquiera se representará por potencias de base 2, tal como:

54 en decimal (54) equivale a 00110110

$$\begin{aligned}
 54 &= 00110110 = 0.2^7 + 0.2^6 + 1.2^5 + 1.2^4 + 0.2^3 + 1.2^2 + 1.2^1 + 0.2^0 \\
 &= 0 + 0 + 32 + 16 + 0 + 4 + 2 + 0 = 54
 \end{aligned}$$

La Tabla 1.7 representa notaciones equivalentes de los cuatro sistemas de numeración comentados anteriormente.

Tabla 1.7. Equivalencias de códigos decimal, binario, octal y hexadecimal

Decimal	Binario	Octal	Hexadecimal
0	0	0	0
1	1	1	1
2	10	2	2
3	11	3	3
4	100	4	4
5	101	5	5
6	110	6	6
7	111	7	7
8	1000	10	8
9	1001	11	9
10	1010	12	A
11	1011	13	B
12	1100	14	C
13	1101	15	D
14	1110	16	E
15	1111	17	F
16	10000	20	10
17	10001	21	11
18	10010	22	12
19	10011	23	13
20	10100	24	14

1.5. DISPOSITIVOS DE ALMACENAMIENTO SECUNDARIO (ALMACENAMIENTO MASIVO)

La memoria secundaria, mediante los dispositivos de almacenamiento secundario, proporciona capacidad de almacenamiento fuera de la UCP y del almacenamiento o memoria principal. El almacenamiento secundario es no volátil y mantiene los datos y programas, incluso cuando se apaga la computadora. Las unidades (*drives*, en inglés), periféricos o dispositivos de almacenamiento secundario son dispositivos periféricos que actúan como medio de soporte para almacenar datos —temporal o permanentemente— que ha de manipular la UCP durante el proceso en curso y que no puede contener la memoria principal.

Las tecnologías de almacenamiento secundario más importantes son **discos magnéticos**, **discos ópticos** y **cintas magnéticas**. El dispositivo de almacenamiento secundario más común es la unidad de disco o disquete, que sirve para alojar los discos. En ella se almacenan y recuperan datos y programas de un disco, transfiriendo los datos entre la memoria secundaria y la memoria principal.

La información almacenada en la memoria central es *volátil* (desaparece cuando se apaga la computadora) y la información almacenada en la memoria auxiliar es *permanente*. Esta información contenida en la memoria secundaria se conserva en unidades de almacenamiento denominadas **archivos** (**ficheros**, *files* en inglés) que pueden ser tan grandes como se desee. Un programa, por ejemplo, se almacena en un archivo y se copia en memoria principal cuando se ejecuta el programa. Se puede almacenar desde un programa, hasta un capítulo de un libro, un inventario de un almacén o un listado de clientes o cualquier otra unidad de información como música, archivos **MP3**, **DivX**, un **correo electrónico**, etc. Los resultados de los programas se pueden guardar como *archivos de datos* y los programas que se escriben se guardan como *archivos de programas*, ambos en la memoria auxiliar. Cualquier tipo de archivo se puede transferir fácilmente desde la memoria auxiliar hasta la central para su proceso posterior.

1.5.1. Discos magnéticos

Los discos son dispositivos formados por componentes electromagnéticos que permiten un acceso rápido a bloques físicos de datos. La información se registra en la superficie del disco y se accede a ella por medio de cabezas de *lectura/escritura* que se mueven sobre la superficie. Los discos magnéticos se clasifican en **disquetes** (*floppy disk*), ya prácticamente en desuso, y **discos duros** (*hard disk*).

Los primeros disquetes, antes del advenimiento del PC eran de 8 pulgadas; posteriormente aparecieron del tamaño de 5 1/4" de 360 KB de capacidad que llegaron a alcanzar 1,2 MB (ya prácticamente en desuso) y los que se fabrican en la actualidad de 3,5" y capacidad de 1,44 Megabytes (2,8 MB, en algunos casos). Los disquetes han sido muy populares, pero hoy día cada vez se utilizan menos, su gran ventaja era su tamaño y que eran transportables de una computadora a otra, además, era relativamente fácil grabar y borrar su información. Los **discos duros** también llamados discos fijos (*hard disk*) se caracterizan por su gran capacidad de almacenamiento (del orden de decenas, centenas y millares de GB, TB, etc.) y porque normalmente se encuentran empotrados en la unidad física de la computadora. Las computadoras grandes utilizan múltiples discos duros ya que ellos requieren gran capacidad de almacenamiento que se mide en Gigabytes o en Terabytes. Es posible ampliar el tamaño de los discos duros de una computadora, bien cambiándolos físicamente por otros de capacidad mayor o bien añadiendo otros a los existentes.

Un disco debe ser formateado antes de ser utilizado. La operación de formateado escribe información en el disco de modo que los datos se pueden escribir y recuperar eficientemente. El proceso de formatear un disquete es análogo al proceso de dibujar líneas en un aparcamiento y la numeración de las correspondientes plazas. Permite que la información se sitúe (plaza de aparcamiento o "parqueo") y se recupere (encontrar su automóvil de modo rápido y seguro). Esto explica por qué un disco tiene menos espacio en el mismo después de que ha sido formateado (al igual que un aparcamiento, ya que las líneas y la numeración ocupan un espacio determinado).

Hoy día se comercializan numerosos discos duros transportables (*removibles*) que se conectan fácilmente mediante los controladores USB que se verán posteriormente y que se comercializan con tamaños de centenares de MB hasta 1 y 2 TB.

1.5.2. Discos ópticos: CD-ROM y DVD

Los **discos ópticos** difieren de los tradicionales discos duros o discos magnéticos en que los primeros utilizan un haz de láser para grabar la información. Son dispositivos de almacenamiento que utilizan la misma tecnología que los dispositivos compactos de audio para almacenar información digital. Por esta razón suelen tener las mismas caracte-

rísticas que los discos de música: muy resistentes al paso del tiempo y con gran capacidad de almacenamiento. Estos discos se suelen utilizar para almacenar información histórica (no va a sufrir modificaciones frecuentes), archivos gráficos complejos, imágenes digitales, etc. Al igual que los disquetes, son transportables y compatibles entre computadoras. Los dos grandes modelos existentes en la actualidad son los discos compactos (CD) y los discos versátiles digitales (DVD).

El CD-ROM (el cederrón)¹³ (Compact Disk-Read Only Memory, Disco compacto - Memoria de solo lectura)

Estos discos son el medio ideal para almacenar información de forma masiva que no necesita ser actualizada con frecuencia (dibujos, fotografías, enciclopedias...). La llegada de estos discos al mercado hizo posible el desarrollo de la *multimedia*, es decir, la capacidad de integrar medios de todo tipo (texto, sonido e imágenes). Permiten almacenar 650 o 700 Megabytes de información. En la actualidad son muy económicos, alrededor de medio euro (medio dólar). Estos discos son de sólo lectura, por lo que sólo se pueden grabar una vez. Estos discos conocidos como CD-R o CD+R son cada día más populares y han sustituido a los disquetes de 3,5".

Existen discos CD que permiten grabación de datos, además de lectura y se conocen como discos CD-RW (CD-Recordable y ReWritable). Desde hace años es posible encontrar en el mercado estos discos ópticos CD en los que se puede leer y escribir información por parte del usuario cuantas veces se deseen. Es el modelo *regrabable*, por excelencia. Este modelo se suele utilizar para realizar copias de seguridad del disco duro o de la información más sensible, al poder actualizarse continuamente. Aunque nació para emplearse en servidores, estaciones de trabajo, etc., hoy día, es un disco que suele utilizarse en computadoras personales de grandes prestaciones. Las unidades lectoras y grabadoras de discos¹⁴ de este tipo, tiene ya precios asequibles y son muchos los usuarios, incluso, domésticos, que incorporan estas unidades a sus equipos informáticos.

DVD (Digital Versatile Disc): Videodisco digital (DVD-+RW, DVD de alta capacidad de almacenamiento: HD DVD y Blu-ray)

Este disco óptico nació en 1995, gracias a un acuerdo entre los grandes fabricantes de electrónica de consumo, estudios de cine y de música (Toshiba, Philips, Hitachi, JVC, etc.). Son dispositivos de alta capacidad de almacenamiento, interactivos y con total compatibilidad con los medios existentes. Tiene además una gran ventaja: su formato sirve tanto para las computadoras como para los dispositivos de electrónica de consumo. El DVD es capaz de almacenar hasta 26 CD con una calidad muy alta y con una capacidad que varía, desde los 4,7 GB del tipo de una cara y una capa hasta los 17 GB del de dos caras y dos capas, o lo que es igual, el equivalente a la capacidad de 7 a 26 CD convencionales. Estas cifras significan que se pueden almacenar en uno de estos discos una película completa en diferentes idiomas e incluso subtítulos.

En la actualidad se pueden encontrar tres formatos de DVD grabables: **DVD-R** (se puede grabar una sola vez); **DVD-RAM** (*reescribible* pero con un funcionamiento similar al disco duro); **DVD-RW** (lectura y escritura, regrabable). Al igual que en el caso de los discos compactos, requieren de unas unidades especiales de lectura y reproducción, así como grabadoras/regrabadoras. Estas últimas se encuentran ya en el mercado, a precios muy asequibles. La mayoría de las computadoras que se comercializan en cualquier gran almacén incluyen de serie una unidad lectora de DVD y grabadora de CD-RW o de DVD, que permiten grabar una y otra vez en los discos de formato RW. Comienza a ser también frecuente encontrar PCs con unidades de grabación de todos los formatos DVD, tales como **DVD-R**, **DVD+R**, **DVD-RW** y **DVD+RW** y, ya son una realidad, los nuevos DVD de alta definición de Toshiba y Blu-ray de Sony de alta capacidad de almacenamiento (15 GB a 50 GB). En abril de 2006 se presentaron los 16 nuevos lectores de DVD de gran capacidad de almacenamiento de Toshiba (HD DVD, de 15 GB a 30 GB) y **Blu-ray** de Sony (de 25 GB a 50 GB), y a finales del primer semestre de 2007, Toshiba presentó su nuevo HD_DVD de 3 capas con lo que llegó a 51 GB y así competir directamente con Sony también en capacidad (Figura 1.6).

Discos duros virtuales

Es un nuevo dispositivo de almacenamiento de información que no reside en la computadora del usuario sino en un espacio virtual residente en un sitio Web de Internet (de tu propia empresa, o de cualquiera otra que ofrezca el servicio).

¹³ La última edición (22.ª, 2001) del Diccionario de la Lengua Española (DRAE) ha incorporado el término **cederrón**.

¹⁴ En Hispanoamérica se conoce también a estas unidades como unidades "quemadoras" de disco, traducción fiel del término anglosajón.



Figura 1.6. Unidad de disco USB (arriba izquierda), unidad de DVD regrabable (arriba derecha), lector de Blu-ray (abajo).

Es una buena opción para el usuario (estudiantes, particulares, profesionales, empresas...) de tipo medio y empresas que utilizan grandes volúmenes de información y que necesitan más espacio y no lo tienen disponible en sus equipos. Este almacenamiento o alojamiento puede ser gratuito o de pago, pero en cualquier forma no deja de ser una interesante oferta para el programador que encuentra un lugar donde situar aplicaciones, archivos, etc., que no puede almacenar en su computadora.

El inconveniente de esta solución es el riesgo que se ha de asumir al depositar información en lugares no controlados por uno mismo. Esta situación plantea la necesidad de un estudio de la privacidad y seguridad que van a tener los datos que deposite en estos discos virtuales. La Tabla 1.8 muestra algunas direcciones de almacenamiento virtual en Internet que en algunos casos son gratuitos.

Tabla 1.8. Algunas direcciones de sitios Web para almacenamiento virtual de datos

Nombre de la empresa	Dirección de Internet
<i>Xdrive</i>	www.xdrive.com
FreeDrive (propiedad de <i>Xdrive</i>)	www.freedrive.com
<i>FreeMailGuide</i>	www.freemailguide.com
<i>Yahoo; Briefcase</i> (necesita registro previo)	briefcase.yahoo.com

Hoy, además de sitios como los referenciados en la tabla anterior, la mayoría de los buscadores de Internet ofrecen una gran capacidad de almacenamiento gratuito, donde se pueden almacenar gran cantidad de datos además de los correos electrónicos, y totalmente gratuitos, y con un programa adecuado se puede también convertir este espacio de correo electrónico en espacio para un disco duro virtual. Los servicios de correo electrónico (webmail) ofrecen capacidad de almacenamiento creciente que pueden ser muy bien utilizados por el internauta¹⁵.

¹⁵ La tendencia en 2008, es aumentar el almacenamiento gratuito que ofrecen los grandes buscadores. Gmail, ofrece en febrero de 2008, la cantidad de 6389 MB, o sea más de 6,2 GB, de almacenamiento gratuito a sus clientes de correo electrónico. Yahoo! ofrece almacenamiento ilimitado y Windows Live Hotmail más de 5 GB.

1.5.3. Discos y memorias Flash USB

Los *chips* de memoria flash, similares a los chips de RAM, son unos chips con una tecnología especial, flash, en los que se puede escribir y borrar rápida y repetidamente, pero al contrario que las memorias RAM, las memorias flash no son volátiles y se puede mantener su contenido sin alimentación eléctrica.

Cámaras digitales, teléfonos celulares (móviles), computadoras portátiles, PDA, y otros dispositivos digitales utilizan memoria flash para almacenar datos que necesitan modificarse en el transcurso del tiempo.

Las memorias flash siguen siendo muy caras aunque el proceso de abaratamiento se ha iniciado en estos últimos años y pronto reemplazarán a discos y chips de memoria tradicionales. Hoy día, finales de 2007, es relativamente fácil encontrar tarjetas de memorias flash o lápices USB (pen drives) de 1 GB a 8 GB por precios muy asequibles (15 a 30 €) y la tendencia es aumentar la cantidad de memoria que almacena y reducción del precio.

Asimismo los discos duros externos con conexiones mediante USB se comercializan con tamaños de memoria de cientos de GB hasta Terabytes (1 y 2 TB son capacidades de unidades de disco externo USB que se encuentran fácilmente en grandes almacenes y tiendas especializadas y también con precios asequibles en torno a 100 y 200 €).

Una memoria *flash*, también comercializada como un disco es un pequeño almacén de memoria móvil de un tamaño algo mayor que un mechero o llavero (por esta razón a veces se les llama *llaveros flash*) y por consiguiente se puede transportar en el bolsillo de una prenda de vestir. Este disco o memoria se puede conectar a cualquier PC de escritorio o portátil que disponga de una conexión USB (véase apartado 1.4.2). Se comercializa por muchos fabricantes¹⁶ y se han convertido en el medio más económico y práctico para llevar archivos de cualquier tipo e incluso hasta programas como copias de seguridad. Los discos duros USB al ser regrabables y de fácil instalación (sólo necesitan enchufarse en un puerto USB) se están constituyendo en el medio idóneo para almacenamiento de información personal y como dispositivo de copia de seguridad.



Figura 1.7. Tarjeta compact flash (izquierda), memoria flash USB (centro) y disco duro (derecha).

1.5.4. Otros dispositivos de Entrada y Salida (E/S)

Los dispositivos de entrada y de salida permiten la comunicación entre las personas y la UCP. Un **dispositivo de entrada** es cualquier dispositivo que permite que una persona envíe información a la computadora. Los dispositivos de entrada, por excelencia, son un teclado y un ratón. Entre otras cosas un ratón se utiliza para apuntar, moverse por la pantalla y elegir una lista de opciones visualizadas en la pantalla. El dispositivo fue bautizado como *ratón* (*mouse* en inglés, jerga muy utilizada también en Latinoamérica) porque se conecta a la computadora por un largo cable y el conjunto se asemeja a un ratón. El ratón típico tiene dos o tres botones, e incluso una pequeña ruedecita que permite desplazarse por menús y similares en la pantalla. El puntero en la pantalla se conoce como cursor o *sprite*. Moviéndose con el ratón de modo que el cursor apunte a una región específica de la pantalla (por ejemplo, un menú de una aplicación) y haciendo clic en el botón del ratón, se puede señalar a la computadora para que realice la orden indicada en la opción del menú. El uso del ratón y de menús facilita dar órdenes a la computadora y es mucho más sencillo que las tediosas órdenes de teclado que siempre se deben memorizar. Algunos dispositivos de entrada, no tan típicos pero cada vez más usuales en las configuraciones de sistemas informáticos son: escáner, lápiz óptico, micrófono y reconocedor de voz.

¹⁶ En febrero de 2006.

Un **dispositivo de salida** es cualquier dispositivo que permite a una computadora pasar información al usuario. El dispositivo de salida por excelencia es la pantalla de presentación, también llamada *monitor* o *terminal*. Otro dispositivo de salida muy usual es la impresora para producir salidas impresas en papel. Al teclado y la pantalla integrados se les suele conocer también como *terminal* o *VDT* (*video display terminal*).

El monitor, conocido también como CRT (*cathode ray tube*) funciona igual que un aparato de televisión. El monitor está controlado por un dispositivo de salida denominado *tarjeta gráfica*. Las tarjetas gráficas envían los datos para ser visualizados en el monitor con un formato que el monitor puede manipular. Las características más importantes del monitor y la tarjeta gráfica son la velocidad de refresco, la resolución y el número de colores soportados. La *velocidad de refresco* es la velocidad a la cual la tarjeta gráfica actualiza la imagen en la pantalla. Una tasa de refresco baja tal como 60 KHz, puede producir fatiga en los ojos ya que la imagen puede parpadear imperceptiblemente. Las tarjetas gráficas usuales presentan tasas de refresco de 70 a 100 MHz. Esta frecuencia elimina el parpadeo y la consiguiente fatiga para los ojos. La *resolución* es el número de puntos por pulgada que se pueden visualizar a lo largo de la pantalla. Un punto (*dot*) en este contexto se conoce como un *píxel* (*picture elemental*). En los monitores clásicos VGA una resolución típica es 640×480 : hay 640 pixels en el sentido horizontal de la pantalla y 480 pixels en el vertical. La tarjeta gráfica almacena la información en la pantalla para cada píxel en su propia memoria. Las tarjetas gráficas que pueden visualizar a resoluciones más altas requieren más memoria. Por ejemplo muchas tarjetas soportan resoluciones que corren desde 800×640 hasta 12.180×1.024 . Tales tarjetas requieren 1 a 4 Mb de memoria. Relacionado directamente con la cantidad de memoria y la resolución es el número de colores que se pueden visualizar. La tarjeta gráfica debe almacenar la información del color para visualizar cada píxel en la pantalla. Para visualizar 256 (2^8) colores, se necesita 1 byte por cada píxel.

Dado que las personas y las computadoras utilizan lenguajes diferentes se requiere algún proceso de traducción. Las interacciones con un teclado, la pantalla o la impresora tienen lugar en el idioma español, el inglés o cualquier otro como el catalán. Eso significa que en la jerga informática cuando se pulsa la letra C (de Carhelejo) en un teclado se produce que una letra C vaya a la pantalla del monitor, o a una impresora y allí se visualice o se imprima como una letra C. Existen diversos códigos de uso frecuente. El código más usual entre computadoras es el ASCII (acrónimo de American Standard Code for Information Interchange) que es un código de siete bits que soporta letras mayúsculas y minúsculas del alfabeto, signos numéricos y de puntuación, y caracteres de control. Cada dispositivo tiene su propio conjunto de códigos pero los códigos construidos para un dispositivo no son necesariamente los mismos códigos construidos para otros dispositivos. Algunos caracteres, especialmente caracteres tales como tabulaciones, avances de línea o de página y retornos de carro son manipulados de modo diferente por dispositivos diferentes e incluso por piezas diferentes de sistemas *software* que corren sobre el mismo dispositivo. Desde la aparición del lenguaje Java y su extensión para aplicaciones en Internet se está haciendo muy popular el código Unicode que facilita la integración de alfabetos de lenguajes muy diversos no sólo los occidentales, sino orientales, árabes, etc.

Nuevos dispositivos de E/S móviles

Los sistemas de transmisión de datos que envían señales a través del aire o del espacio sin ninguna atadura física se han vuelto una alternativa fiable a los canales cableados tradicionales tales como el cable de cobre, cable coaxial o de fibra óptica. Hoy en programación se utilizan como dispositivos de E/S, teléfonos inteligentes (*smartphones*), asistentes digitales personales, PDA y redes de datos móviles.

Los teléfonos móviles (celulares) son dispositivos que transmiten voz o datos (últimamente también imágenes y sonidos) que utilizan ondas radio para comunicarse con antenas de radios situados en celdas (áreas geográficas adyacentes) que a su vez se comunican con otras celdas hasta llegar a su destino, donde se transmiten al teléfono receptor o al servidor de la computadora al que está conectado. Los nuevos modelos de teléfonos digitales pueden manejar correo voz, correo electrónico y faxes, almacenan direcciones, acceden a redes privadas corporativas y a información de Internet. Los **teléfonos inteligentes** vienen equipados con *software* de navegación Web que permite a estos dispositivos acceder a páginas Web cuyos formatos han sido adaptados al tamaño de sus pantallas.

Los asistentes personales digitales (PDA) son pequeñas computadoras de mano capaces de realizar transmisiones de comunicaciones digitales. Pueden incorporar¹⁷ telecomunicaciones inalámbricas y *software* de organización del trabajo de oficina o para ayuda al estudio. Nokia, Palm, HP, Microsoft son algunos de los fabricantes que construyen este tipo de dispositivos. Los teléfonos móviles o celulares y los PDAs pueden venir incorporados con tecnologías GPRS o tecnología UMTS/CDMA. Las tecnologías GPRS conocidas como generación 2.5 per-

¹⁷ Este es el caso del PDA del fabricante español Airis que comercializa a un coste asequible, un teléfono/PDA.

miten velocidades de transmisión de 50 a 100 Kbps, similar y un poco mayor a la velocidad de la red de telefonía básica, RTB. Los teléfonos UMTS/CDMA que ya se comercializan en Europa¹⁸ y también en América y Asia, se conocen como teléfonos de 3.^a generación (3G), y permiten velocidades de transmisión hasta 1 o 2 Mbps, igual cantidad que las telefonías digitales ADSL.



Figura 1.8. Blackberry (izquierda), Palm Treo (centro) y HP iPAQ hw6500 (derecha).

1.6. CONECTORES DE DISPOSITIVOS DE E/S

Los dispositivos de E/S no se pueden conectar directamente a la UCP y la memoria, dada su diferente naturaleza. Los dispositivos de E/S son dispositivos electromecánicos, magnéticos u ópticos que además funcionan a diferentes velocidades, la UCP y la memoria son dispositivos electrónicos. Por otra parte los dispositivos de E/S operan a una velocidad mucho más lenta que la UCP/memoria. Se requiere por consiguiente de un dispositivo intermedio o adaptador denominado **interfaz** o **controlador**. Existe un controlador específico para cada dispositivo de entrada/salida que puede ser de *software* o de *hardware*. Los controladores de *hardware* más utilizados presentan al exterior conectores donde se enchufan o conectan los diferentes dispositivos. Cada computadora tiene un número determinado de conectores estándar incorporados y que se localizan fácilmente en el exterior de su chasis. Los sistemas operativos modernos como Windows XP reconocen automáticamente los dispositivos de E/S tan pronto se conectan a la computadora. Si no es así necesitará cargar en memoria un programa de *software* denominado controlador del dispositivo correspondiente con el objetivo de que el sistema operativo reconozca al citado dispositivo. Los conectores más comunes son: puertos serie y paralelo, buses USB y *firewire*.

1.6.1. Puertos serie y paralelo

El PC está equipado con puertos serie y paralelo. El puerto serie (como mínimo suele tener dos) es un conector macho de la parte trasera o lateral del PC con 9 o 25 clavijas, aunque sólo suelen utilizarse 3 o 4 para la transmisión en serie. El puerto paralelo también se denomina puerto de impresora, ya que es donde solía conectarse la impresora

¹⁸ Ya comienza a extenderse, al menos en el ámbito empresarial, las tarjetas digitales del tipo PCMCIA, 2.5G/3G que son tarjetas módem 2G/3G con una memoria SIM y número teléfono móvil (celular) incorporado y que enchufadas a una computadora portátil permiten conexiones a Internet a velocidad UMTS y en aquellas zonas geográficas donde no exista cobertura, automáticamente se conecta a velocidad 2.5 G (GPRS) que tiene mayor cobertura en el resto del territorio. En España desde el mes de julio de 2004, tanto Vodafone como Telefónica Móviles ofrecen estas soluciones.

hasta que aparecieron los conectores USB. El conector de la impresora de la parte trasera del PC es un conector hembra de 25 clavijas. Los puertos se llaman también **COM1**, **COM2** y **LPT** conocidos por nombres de dispositivos lógicos que el programa de inicio del PC automáticamente asigna a estos dispositivos durante el inicio, por ejemplo A:, C:, E:, CON, PRN y KBD son nombres lógicos.

1.6.2. USB

USB son las siglas de Universal Serial Bus (Bus serie universal) y corresponden a un *bus* estándar de E/S que desarrollaron originalmente varias empresas, entre ellas Compaq, Digital, IBM, Intel, Microsoft, NEC y Northern Telecom¹⁹. La importancia del *bus* USB es que es un bus de E/S serie de precio asequible con una especificación práctica, lo que significa que cualquiera puede producir productos USB sin tener que pagar ninguna licencia. Sin duda, el *bus* USB es la innovación más importante y de éxito del mundo PC en muchos años. Es un bus de expansión que permite conectar una gran cantidad de equipamiento al PC.

El objetivo del USB conseguido es reunir las diferentes conexiones del teclado, el ratón, el escáner, el joystick, la cámara digital, impresora, disco duro, etc., en un *bus* compartido conectado a través de un tipo de conector común. Otra gran ventaja es también su compatibilidad con computadoras Macintosh. Existen dos versiones: **USB 1.1** cuya velocidad de transferencia está limitada a un máximo de 12 Mbps; **USB 2.0** puede transmitir hasta 40 Mbps y se utiliza en todos los PC modernos. La versión 2.0 es compatible descendente; es decir, un dispositivo con un conector USB 2.0 es compatible con los conectores 1.1 y no siempre sucede igual al revés. Otra gran ventaja es que ya se fabrican distribuidores (*hubs*) que permiten conectar numerosos dispositivos USB a un único *bus* USB. Con independencia de la conexión de distribuidores USB, ya es frecuente que tanto los PC de escritorio como los portátiles vengan de fábrica con un número variable de 2 a 8 e incluso 10 puertos USB, normalmente el estándar 2.0.

1.6.3. Bus IEEE Firewire – 1394

El *bus* IEEE 1394 (*firewire*) es una nueva interfaz SCSI (un *bus* antiguo pero avanzado utilizado para discos duros, unidades de CD-ROM, escáneres y unidades de cinta). Es un *bus* serie de alta velocidad con una velocidad de transferencia máxima de 400 Mbps patentado por Apple. Es una interfaz estándar de bus serie para computadoras personales (y vídeo/audio digital). **IEEE 1394** ha sido adoptado como la interfaz de conexiones estándar **HANA** (*High Definition Audio-Video Network Alliance*) para comunicación y control de componentes audiovisuales. *Firewire* está también disponible en versiones inalámbricas (*wireless*), fibra óptica y cable coaxial. Las computadoras Apple y Sony suelen venir con puertos *firewire*, y ya comienza a ser usual que los PC incluyan al menos un puerto *firewire*. Las actuales videocámaras digitales y otros dispositivos de audio e imagen suelen incorporar conectores *firewire*.



Figura 1.9. Conectores USB (izquierda) y conector Firewire (derecha).

¹⁹ En el sitio www.usb.org y en el forum “USB Implementers Forum” puede encontrar historia y características del bus USB.

1.7. REDES, WEB Y WEB 2.0

Hoy día las computadoras autónomas (*standalone*) prácticamente no se utilizan (excepción hecha del hogar) y están siendo reemplazadas hasta en los hogares y en las pequeñas empresas, por redes de computadoras. Una red es un conjunto de computadoras conectadas entre sí para compartir recursos. Al contrario que una gran computadora que es una única computadora compartida por muchos usuarios, una red (*network*) consta de muchas computadoras que comparten recursos.

Las computadoras modernas necesitan comunicarse con otras computadoras. Si la computadora se conecta con una *tarjeta de red* se puede conectar a una red de datos locales (*red de área local*). De este modo se puede acceder y compartir a cada una de las memorias de disco y otros dispositivos de entrada y salida. Si la computadora tiene un *módem*, se puede comunicar con computadoras distantes. Se pueden conectar a una red de datos o *enviar correo electrónico* a través de las redes corporativas Intranet/Extranet o la propia red Internet. También es posible enviar y recibir mensajes de fax.

El uso de múltiples computadoras enlazadas por una red de comunicaciones para distribuir el proceso se denomina proceso distribuido en contraste con el proceso centralizado en el cual todo el proceso se realiza por una computadora central. De esta forma los sistemas de computadoras también se clasifican en **sistemas distribuidos** y **sistemas centralizados**.

Las redes se pueden clasificar en varias categorías siendo las más conocidas las redes de área local (LAN, Local Area Network) y las redes área amplia o ancha WAN (Wide Area Network). Una Red de Área Local permite a muchas computadoras acceder a recursos compartidos de una computadora más potente denominada servidor. Una WAN es una red que enlaza muchas computadoras personales y redes de área local en una zona geográfica amplia. La red WAN más conocida y popular en la actualidad es la red Internet que está soportada por la World Wide Web.

Una de las posibilidades más interesantes de las computadoras es la comunicación entre ellas cuando se encuentran en sitios separados físicamente y se encuentran enlazadas por vía telefónica. Estas computadoras se conectan en redes LAN (Red de Área Local) y WAN (Red de Área Ancha), aunque hoy día las redes más implantadas son las redes que se conectan con tecnología Internet y por tanto conexión a la Red Internet. Estas redes son *Intranet* y *Extranet* y se conocen como redes corporativas ya que enlazan computadoras de los empleados de las empresas. Las instalaciones de las comunicaciones requieren de líneas telefónicas analógicas o digitales y de modems.

Los sistemas distribuidos realizan el proceso de sus operaciones de varias formas siendo las más conocidas cliente-servidor e igual-a-igual (*peer-to-peer*, *P2P*).

Compartición de recursos

Uno de los usos más extendidos de la red es permitir a diferentes computadoras compartir recursos tales como sistemas de archivos, impresoras, escáneres o discos DVD. Estas computadoras normalmente se conectan en una relación denominada **cliente-servidor** (Figura 1.10). El servidor posee los recursos que se quieren compartir. Los clientes conectados vía un concentrador (*hub*) o una conexión *ethernet* comparten el uso de estos recursos. El usuario de una

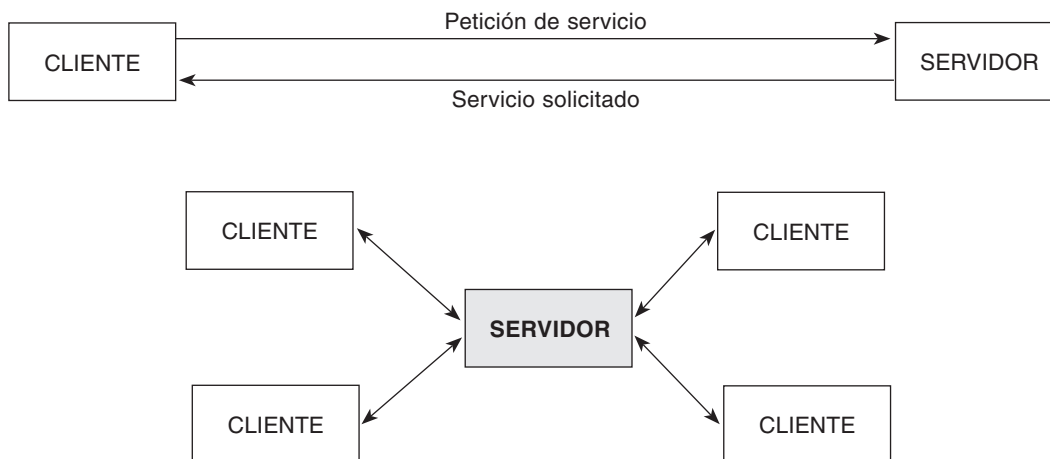


Figura 1.10. Sistema de computadoras Cliente/Servidor.

máquina cliente puede imprimir documentos o acceder a archivos como si los dispositivos realmente estuvieran físicamente conectados a la máquina local. Esto puede dar la ilusión de que realmente se tienen más recursos de los que realmente existen, así como un entorno de programación uniforme, independiente de la máquina que realmente se utilice.

El **sistema cliente-servidor** es el más popular en computación. El sistema divide el procesamiento de las tareas entre las computadoras “cliente” y las computadoras “servidor” que a su vez están conectadas en red. A cada máquina se le asignan funciones adecuadas a sus características. El **cliente** es el usuario final o punto de entrada a la red y normalmente en una computadora personal de escritorio o portátil, o una estación de trabajo. El usuario, normalmente interactúa directamente sólo con la parte cliente del sistema, normalmente, para entrada o recuperación de información y uso de aplicaciones para análisis y cálculos posteriores.

El **servidor** proporciona recursos y servicios a otras computadoras de la red (los clientes). El servidor puede ser desde una gran computadora a otra computadora de escritorio pero especializada para esta finalidad y mucho más potente. Los servidores almacenan y procesan los datos compartidos y también realizan las funciones no visibles, de segundo plano (*back-end*), a los usuarios, tales como actividades de gestión de red, implementación de bases de datos, etc. La Figura 1.10 muestra un sistema cliente/servidor. La red Internet es el sistema cliente/servidor más popular.

1.7.1. Redes P2P, igual-a-igual (*peer-to-peer*, P2P)

Otra forma de sistema distribuido es la computación **P2P**²⁰ (*peer-to-peer*) que es un sistema que enlaza las computadoras vía Internet o redes privadas de modo que pueden compartir tareas de proceso. El modelo P2P se diferencia del modelo de red cliente/servidor en que la potencia de proceso reside sólo en las computadoras individuales de modo que trabajan juntos colaborando entre sí, pero sin un servidor o cualquier otra computadora que los controle. Los sistemas P2P utilizan espacio de disco o potencia de proceso del PC no utilizado por los sistemas en red. Estos sistemas P2P se utilizan hoy con gran profusión en ambientes científicos y de investigación, así como para descargas de música por Internet.

1.7.2. Aplicaciones de las redes de comunicaciones

En el interior de la computadora los diferentes componentes de *hardware* se comunican entre sí utilizando el *bus* interno. Hoy día es práctica común que las computadoras se comuniquen unas con otras compartiendo recursos e información. Esta actividad es posible a través del uso de redes, con cables físicos (normalmente teléfonos alámbricos), junto con transmisiones electrónicas, sin cables (inalámbricas) mediante teléfonos móviles o celulares, redes inalámbricas o tecnologías Bluetooth.

Existen muchos tipos de redes. Una **red de área local (LAN, local area network)** normalmente une decenas y a veces centenares de computadoras en una pequeña empresa u organismo público. Una **red global**, tal como Internet, que se expande a distancias mucho mayores y conecta centenares o millares de máquinas que, a su vez, se unen a redes más pequeñas a través de computadoras pasarela (*gateway*). Una computadora pasarela (*gateway*) es un puente entre una red tal como Internet en un lado y una red de área local en el otro lado. La computadora también suele actuar como un **cortafuegos (firewall)** cuyo propósito es mantener las transmisiones ilegales, no deseadas o peligrosas fuera del entorno local. Estas redes se suelen conocer normalmente como redes **Intranet** y en realidad son redes corporativas o institucionales que utilizan tecnología Internet y que por consiguiente pueden enlazarse con otras redes de compañías socias, clientes, amigas, etc., y todo tipo de posibles clientes personales e institucionales sin necesidad de que estos a su vez formen una red.

Otro uso típico de redes es la comunicación. El correo electrónico (*e-mail*) se ha convertido en un medio muy popular para enviar cartas y documentos de todo tipo así como archivos a amigos, clientes, socios, etc. La World Wide Web está proporcionando nuevas oportunidades comerciales y profesionales tanto a usuarios aislados como a usuarios pertenecientes a entidades y empresas. Las redes han cambiado también los conceptos y hábitos de los lugares de trabajo y el trabajo en sí mismo. Muchos estudiantes y profesionales utilizan las transmisiones de las redes entre el hogar y la oficina o entre dos oficinas de modo que puedan acceder a la información que necesitan siempre

²⁰ Los sistemas P2P se hicieron muy populares y llegaron al gran público cuando un estudiante estadounidense, Shawn Fanning, inventó el sistema Napster, un sistema que permite descargas de música entre computadoras personales sin intervención de ningún servidor central.

que lo necesiten, y de hecho desde el lugar que ellos decidan siempre que exista una línea telefónica o un teléfono móvil (celular).

Otro concepto importante es la *informática distribuida*. Las redes se utilizan también para permitir que las computadoras se comuniquen entre sí. La complejidad de muchos problemas actuales requiere el uso de reservas de computación. Esto se puede conseguir por sincronización de los esfuerzos de múltiples computadoras, trabajando todas en paralelo en componentes independientes de un problema. Un sistema distribuido grande puede hacer uso de centenares de computadoras.

1.7.3. Módem

El módem es un dispositivo periférico que permite intercambiar información entre computadoras a través de una línea telefónica. El **módem** es un acrónimo de **Modulador-Demodulador**, y es un dispositivo que transforma las señales digitales de la computadora en señales eléctricas analógicas telefónicas y viceversa, con lo que es posible transmitir y recibir información a través de la línea telefónica.

El módem convierte una señal analógica en señal digital, y viceversa.

Los modems permiten además de las conexiones entre computadoras, envío y recepción de faxes, acceso a Internet, etc. Una de las características importantes de un módem es su velocidad; cifras usuales son 56 kilobaudios (1 **baudio** es 1 bit por segundo, *bps*; *Kbps* son 1.000 baudios).

Los modems pueden ser de tres tipos: **Interno** (es una tarjeta que se conecta a la placa base internamente); **Externo** (es un dispositivo que se conecta externamente a la computadora a través de puertos COM, USB, etc.); **PC-Card**, son modems del tipo tarjeta de crédito, que sirven para la conexión a las computadoras portátiles.

Además de los modems analógicos es posible la conexión con Internet y las redes corporativas de las compañías mediante la Red Digital de Sistemas Integrados (**RDSI**, **ISDN**, en inglés) que permite la conexión a 128 Kbps, disponiendo de dos líneas telefónicas, cada una de ellas a 64 Kbps (hoy día ya es poco utilizada). En la actualidad se está implantando a gran velocidad la tecnología digital **ADSL** que permite la conexión a Internet a velocidad superior a la red RDSI, 256 Kbps a 1 a 8 Mbps; son velocidades típicas según sea para “*subir*” datos a la *Red* o para “*bajar*”, respectivamente. Estas cifras suelen darse para accesos personales, ya que en accesos profesionales se pueden alcanzar velocidades de hasta 20-40 Mbps, e incluso superior.

1.7.4. Internet y la World Wide Web

Internet, conocida también como la Red de Redes, se basa en la tecnología Cliente/Servidor. Las personas que utilizan la Red controlan sus tareas mediante aplicaciones Web tal como *software* de navegador. Todos los datos incluyendo mensajes de *correo-e* y las páginas Web se almacenan en servidores. Un cliente (usuario) utiliza Internet para solicitar información de un servidor Web determinado situado en una computadora lejana; el servidor envía la información solicitada al cliente vía la red Internet.

Las plataformas cliente incluyen PC y otras computadoras pero también un amplio conjunto de dispositivos electrónicos de mano (*handheld*) tales como PDA, teléfonos móviles, consolas de juegos, etc., que acceden a Internet de modo inalámbrico (sin cables) a través de señales radio.

La World Wide Web (**WWW**) o simplemente la Web fue creada en 1989 por Bernards Lee en el **CERN** (*European Laboratory for Particles Physics*) aunque su difusión masiva comenzó en 1993 como medio de comunicación universal. La Web es un sistema de estándares aceptados universalmente para almacenamiento, recuperación, formateado y visualización de información, utilizando una arquitectura cliente/servidor. Se puede utilizar la Web para enviar, visualizar, recuperar y buscar información o crear una página Web. La Web combina texto, hipermedia, sonidos y gráficos, utilizando interfaces gráficas de usuario para una visualización fácil.

Para acceder a la Web se necesita un programa denominado navegador Web (*browser*). Un navegador²¹ es una interfaz gráfica de usuario que permite “**navegar**” a través de la Web. Se utiliza el navegador para visualizar textos,

²¹ El navegador más utilizado en la actualidad es *Explorer* de Microsoft, aunque *Firefox* alcanzaba ya un 10% del mercado. En su día fueron muy populares *Netscape* y *Mosaic*.

gráficos y sonidos de un documento Web y activar los enlaces (*links*) o conexiones a otros documentos. Cuando se hace clic (con el ratón) en un enlace a otro documento se produce la transferencia de ese documento situado en otra computadora a su propia computadora.

La World Wide Web está constituida por millones de documentos enlazados entre sí, denominados páginas Web. Una página Web, normalmente, está construida por texto, imágenes, audio y vídeo, al estilo de la página de un libro. Una colección de páginas relacionadas, almacenadas en la misma computadora, se denomina **sitio Web** (*Web site*). Un sitio Web está organizado alrededor de una página inicial (*home page*) que sirve como página de entrada y punto de enlace a otras páginas del sitio. En el párrafo siguiente se describe cómo se construye una página Web. Cada página Web tiene una dirección única, conocida como **URL** (*Uniform Resource Locator*). Por ejemplo, la URL de la página inicial de este libro es: `www.mhe.es/joyanes`.

La Web se basa en un lenguaje estándar de hipertexto denominado **HTML** (*Hypertext Markup Language*) que da formatos a documentos e incorpora enlaces dinámicos a otros documentos almacenados en la misma computadora o en computadoras remotas. El navegador Web está programado de acuerdo al estándar citado. Los documentos HTML, cuando, ya se han situado en Internet, se conocen como páginas Web y el conjunto de páginas Web pertenecientes a una misma entidad (empresa, departamento, usuario individual) se conoce como **sitio Web** (*Website*). En los últimos años ha aparecido un nuevo lenguaje de marcación para formatos, heredero de HTML, y que se está convirtiendo en estándar universal, es el lenguaje **XML**.

Otros servicios que proporciona la Web y ya muy populares para su uso en el mundo de la programación son: el **correo electrónico** y la **mensajería instantánea**. El correo electrónico (*e-mail*) utiliza protocolos específicos para el intercambio de mensajes: **SMTP** (*Simple Mail Transfer Protocol*), **POP** (*Post Office Protocol*) e **IMAP** (*Internet Message Action Protocol*). La mensajería instantánea o *chat* que permite el diálogo en línea simultánea entre dos o más personas, y cuya organización y estructura han sido trasladadas a los teléfonos celulares donde también se puede realizar este tipo de comunicaciones con mensajes conocidos como “cortos” **SMS** (*short message*) o **MMS** (*multimedia message*).

Web 2.0

Este término, ya muy popular, alude a una nueva versión o generación de la Web basada en tecnologías tales como el lenguaje AJAX, los agregadores de noticias RSS, *blogs*, *podcasting*, redes sociales, interfaces de programación de aplicaciones Web (APIs), etc. En esencia, la Web 2.0, cuyo nombre data de 2004, fue empleado por primera vez por Tim O’Reilly, editor de la editorial O’Reilly, ha dado lugar a una Web más participativa y colaborativa, donde el usuario ha dejado de ser un actor pasivo para convertirse en un actor activo y participativo en el uso y desarrollo de aplicaciones Web.



Figura 1.11. Elementos de la siguiente generación de la web. (Fuente: <http://web2.wsj2.com>.)

1.8. EL SOFTWARE (LOS PROGRAMAS)

El *software* de una computadora es un conjunto de instrucciones de programa detalladas que controlan y coordinan los componentes *hardware* de una computadora y controlan las operaciones de un sistema informático. El auge de las computadoras el siglo pasado y en el actual siglo XXI, se debe esencialmente al desarrollo de sucesivas generaciones de *software* potentes y cada vez más *amistosas* (“fáciles de utilizar”).

Las operaciones que debe realizar el *hardware* son especificadas por una lista de instrucciones, llamadas **programas**, o *software*. Un programa de *software* es un conjunto de **sentencias** o **instrucciones** a la computadora. El proceso de escritura o codificación de un programa se denomina **programación** y las personas que se especializan en esta actividad se denominan **programadores**. Existen dos tipos importantes de *software*: *software* del sistema y *software* de aplicaciones. Cada tipo realiza una función diferente.

El **software del sistema** es un conjunto generalizado de programas que gestiona los recursos de la computadora, tal como el procesador central, enlaces de comunicaciones y dispositivos periféricos. Los programadores que escriben *software* del sistema se llaman **programadores de sistemas**. El **software de aplicaciones** es el conjunto de programas escritos por empresas o usuarios individuales o en equipo y que instruyen a la computadora para que ejecute una tarea específica. Los programadores que escriben *software* de aplicaciones se llaman **programadores de aplicaciones**.

Los dos tipos de *software* están relacionados entre sí, de modo que los usuarios y los programadores pueden hacer así un uso eficiente de la computadora. En la Figura 1.12 se muestra una vista organizacional de una computadora donde se ven los diferentes tipos de *software* a modo de capas de la computadora desde su interior (el *hardware*) hasta su exterior (usuario). Las diferentes capas funcionan gracias a las instrucciones específicas (instrucciones máquina) que forman parte del *software* del sistema y llegan al *software* de aplicación, programado por los programadores de aplicaciones, que es utilizado por el usuario que no requiere ser un especialista.



Figura 1.12. Relación entre programas de aplicación y programas del sistema.

1.8.1. Software del sistema

El *software* del sistema coordina las diferentes partes de un sistema de computadora y conecta e interactúa entre el *software* de aplicación y el *hardware* de la computadora. Otro tipo de *software* del sistema que gestiona, controla las actividades de la computadora y realiza tareas de proceso comunes, se denomina *utility* o **utilidades** (en algunas partes de Latinoamérica, **utilerías**). El *software* del sistema que gestiona y controla las actividades de la computadora se denomina **sistema operativo**. Otro *software* del sistema son los **programas traductores** o de traducción de lenguajes de computadora que convierten los lenguajes de programación, entendibles por los programadores, en lenguaje máquina que entienden las computadoras.

El **software del sistema** es el conjunto de programas indispensables para que la máquina funcione; se denominan también *programas del sistema*. Estos programas son, básicamente, *el sistema operativo*, *los editores de texto*, *los compiladores/intérpretes* (lenguajes de programación) y *los programas de utilidad*.

1.8.2. Software de aplicación

El *software* de aplicación tiene como función principal asistir y ayudar a un usuario de una computadora para ejecutar tareas específicas. Los programas de aplicación se pueden desarrollar con diferentes lenguajes y herramientas de *software*. Por ejemplo, una aplicación de procesamiento de textos (*word processing*) tal como Word o Word Perfect que ayuda a crear documentos, una hoja de cálculo tal como Lotus 1-2-3 o Excel que ayudan a automatizar tareas tediosas o repetitivas de cálculos matemáticos o estadísticos, a generar diagramas o gráficos, presentaciones visuales como PowerPoint, o a crear bases de datos como Access u Oracle que ayudan a crear archivos y registros de datos.

Los usuarios, normalmente, compran el *software* de aplicaciones en discos CD o DVD (antiguamente en disquetes) o los descargan (*bajan*) de la Red Internet y han de instalar el *software* copiando los programas correspondientes de los discos en el disco duro de la computadora. Cuando compre estos programas asegúrese de que son compatibles con su computadora y con su sistema operativo. Existe una gran diversidad de programas de aplicación para todo tipo de actividades tanto de modo personal, como de negocios, navegación y manipulación en Internet, gráficos y presentaciones visuales, etc.

Los *lenguajes de programación* sirven para escribir programas que permitan la comunicación usuario/máquina. Unos programas especiales llamados *traductores* (**compiladores** o **intérpretes**) convierten las instrucciones escritas en lenguajes de programación en instrucciones escritas en lenguajes máquina (0 y 1, *bits*) que ésta pueda entender.

Los *programas de utilidad*²² facilitan el uso de la computadora. Un buen ejemplo es un *editor de textos* que permite la escritura y edición de documentos. Este libro ha sido escrito en un editor de textos o *procesador de palabras* (“**word procesor**”).

Los programas que realizan tareas concretas, nóminas, contabilidad, análisis estadístico, etc., es decir, los programas que podrá escribir en C, se denominan *programas de aplicación*. A lo largo del libro se verán pequeños programas de aplicación que muestran los principios de una buena programación de computadora.

Se debe diferenciar entre el acto de crear un programa y la acción de la computadora cuando ejecuta las instrucciones del programa. La creación de un programa se hace inicialmente en papel y a continuación se introduce en la computadora y se convierte en lenguaje entendible por la computadora. La ejecución de un programa requiere una aplicación de una entrada (*datos*) al programa y la obtención de una salida (*resultados*). La entrada puede tener una variedad de formas, tales como números o caracteres alfabéticos. La salida puede también tener formas, tales como datos numéricos o caracteres, señales para controlar equipos o robots, etc. (Figura 1.13).

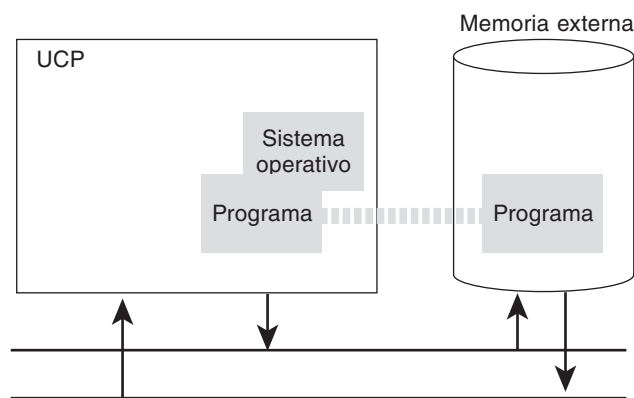


Figura 1.13. Ejecución de un programa.

²² *Utility*: programa de utilidad o utilitería.

1.8.3. Sistema operativo

Un **sistema operativo SO** (*Operating System, OS*) es tal vez la parte más importante del software del sistema y es el software que controla y gestiona los recursos de la computadora. En la práctica el sistema operativo es la colección de programas de computadora que controla la interacción del usuario y el *hardware* de la computadora. El sistema operativo es el administrador principal de la computadora, y por ello a veces se la compara con el director de una orquesta ya que este software es el responsable de dirigir todas las operaciones de la computadora y gestionar todos sus recursos.

El sistema operativo asigna recursos, planifica el uso de recursos y tareas de la computadora, y monitoriza las actividades del sistema informático. Estos recursos incluyen memoria, dispositivos de E/S (Entrada/Salida), y la UCP (Unidad Central de Proceso). El sistema operativo proporciona servicios tales como asignar memoria a un programa y manipulación del control de los dispositivos de E/S tales como el monitor, el teclado o las unidades de disco. La Tabla 1.9 muestra algunos de los sistemas operativos más populares utilizados en enseñanza y en informática profesional.

Tabla 1.9. Sistemas operativos más utilizados en educación y en la empresa

<i>Sistema operativo</i>	<i>Características</i>
Windows Vista	Nuevo sistema operativo de Microsoft presentado a comienzos del año 2007.
Windows XP	Sistema operativo más utilizado en la actualidad, tanto en el campo de la enseñanza, como en la industria y negocios. Su fabricante es Microsoft.
Windows 98/ME/2000	Versiones anteriores de Windows pero que todavía hoy son muy utilizados.
UNIX	Sistema operativo abierto, escrito en C y todavía muy utilizado en el campo profesional.
Linux	Sistema operativo de software abierto, gratuito y de libre distribución, similar a UNIX, y una gran alternativa a Windows. Muy utilizado actualmente en servidores de aplicaciones para Internet.
Mac OS	Sistema operativo de las computadoras Apple Macintosh.
DOS y OS/2	Sistemas operativos creados por Microsoft e IBM respectivamente, ya poco utilizados pero que han sido la base de los actuales sistemas operativos.
CP/M	Sistema operativo de 8 bits para las primeras microcomputadoras nacidas en la década de los setenta.
Symbian	Sistema operativo para teléfonos móviles apoyado fundamentalmente por el fabricante de teléfonos celulares Nokia.
PalmOS	Sistema operativo para agendas digitales, PDA; del fabricante Palm.
Windows Mobile, CE	Sistema operativo para teléfonos móviles con arquitectura y apariencias similares a Windows XP. Las últimas versiones son: 5.0 y 6.0.

Cuando un usuario interactúa con una computadora, la interacción está controlada por el sistema operativo. Un usuario se comunica con un sistema operativo a través de una interfaz de usuario de ese sistema operativo. Los sistemas operativos modernos utilizan una **interfaz gráfica de usuario, IGU** (*Graphical User Interface, GUI*) que hace uso masivo de iconos, botones, barras y cuadros de diálogo para realizar tareas que se controlan por el teclado o el ratón (mouse), entre otros dispositivos.

Normalmente el sistema operativo se almacena de modo permanente en un chip de memoria de sólo lectura (ROM), de modo que esté disponible tan pronto la computadora se pone en marcha (“se enciende” o “se prende”). Otra parte del sistema operativo puede residir en disco, que se almacena en memoria RAM en la inicialización del sistema por primera vez en una operación que se llama *carga* del sistema (*booting*).

El sistema operativo dirige las operaciones globales de la computadora, instruye a la computadora para ejecutar otros programas y controla el almacenamiento y recuperación de archivos (programas y datos) de cintas y discos. Gracias al sistema operativo es posible que el programador pueda introducir y grabar nuevos programas, así como instruir a la computadora para que los ejecute. Los sistemas operativos pueden ser: *monousuarios* (un solo usuario) y *multiusuarios*, o tiempo compartido (diferentes usuarios), atendiendo al número de usuarios y *monocarga* (una sola tarea) o *multitarea* (múltiples tareas) según las tareas (procesos) que puede realizar simultáneamente.

Windows Vista

El 30 de enero de 2007, Microsoft presentó a nivel mundial su nuevo sistema operativo **Windows Vista**. Esta nueva versión en la que Microsoft llevaba trabajando desde hacía cinco años, en que presentó su hasta ahora, última versión, **Windows XP**, es un avance significativo en la nueva generación de sistemas operativos que se utilizarán en la próxima década.

Windows Vista contiene numerosas características nuevas y muchas otras actualizadas, algunas de las cuales son: una interfaz gráfica de usuario muy amigable, herramientas de creación de multimedia, potentes herramientas de comunicación entre computadoras, etc. También ha incluido programas que hasta el momento de su lanzamiento se comercializaban independientemente tales como programas de reproducción de música, vídeo, accesos a Internet, etc. Es de destacar que Vista ha mejorado notablemente la seguridad en el sistema operativo, ya que Windows XP y sus predecesores han sido muy vulnerables a virus, malware, y otros ataques a la seguridad del sistema y del usuario.

Existen cinco versiones comerciales: *Home Basic*, *Home Premium*, *Business*, *Ultimate* y *Enterprise*. Los requisitos que debe tener su computadora dependerá de la versión elegida y variará desde la más básica, recomendada para usuarios domésticos (512 MB de RAM mínima, procesador de 32 bits (x86) o de 64 bits (x64) a 1 GHz, 15 GB de espacio disponible en el disco duro, etc.) a *Ultimate* que incorpora todas las funcionalidades y ventajas contenidas en las demás versiones (ya se requiere al menos 1 GB de memoria, mayor capacidad de disco duro, etc.). A nivel de empresas y grandes corporaciones se recomienda *Enterprise*, diseñada para reducir los riesgos de seguridad y los enormes costes de este tipo de infraestructuras.

Tipos de sistemas operativos

Las diferentes características especializadas del sistema operativo permiten a las computadoras manejar muchas tareas diferentes, así como múltiples usuarios de modo simultáneo o en paralelo, bien de modo secuencial. En función de sus características específicas los sistemas operativos se pueden clasificar en varios grupos.

1.8.3.1. Multiprogramación/Multitarea

La multiprogramación permite a múltiples programas compartir recursos de un sistema de computadora en cualquier momento a través del uso concurrente una UCP. Sólo un programa utiliza realmente la UCP en cualquier momento dado, sin embargo las necesidades de entrada/salida pueden ser atendidas en el mismo momento. Dos o más programas están activos al mismo tiempo, pero no utilizan los recursos de la computadora simultáneamente. Con multiprogramación, un grupo de programas se ejecutan alternativamente y se alternan en el uso del procesador. Cuando se utiliza un sistema operativo de un único usuario, la multiprogramación toma el nombre de **multitarea**.

Multiprogramación

Método de ejecución de dos o más programas concurrentemente utilizando la misma computadora. La UCP ejecuta sólo un programa pero puede atender los servicios de entrada/salida de los otros al mismo tiempo.

1.8.3.2. Tiempo compartido (múltiples usuarios, time sharing)

Un sistema operativo multiusuario es un sistema operativo que tiene la capacidad de permitir que muchos usuarios compartan simultáneamente los recursos de proceso de la computadora. Centenas o millares de usuarios se pueden conectar a la computadora que asigna un tiempo de computador a cada usuario, de modo que a medida que se libera la tarea de un usuario, se realiza la tarea del siguiente, y así sucesivamente. Dada la alta velocidad de transferencia de las operaciones, la sensación es de que todos los usuarios están conectados simultáneamente a la UCP con cada usuario recibiendo únicamente un tiempo de máquina.

1.8.3.3. Multiproceso

Un sistema operativo trabaja en multiproceso cuando puede enlazar dos o más UCP para trabajar en paralelo en un único sistema de computadora. El sistema operativo puede asignar múltiples UCP para ejecutar diferentes instrucciones del mismo programa o de programas diferentes simultáneamente, dividiendo el trabajo entre las diferentes UCP.

La multiprogramación utiliza proceso concurrente con una UCP; el multiproceso utiliza proceso simultáneo con múltiples UCP.

1.9. LENGUAJES DE PROGRAMACIÓN

Como se ha visto en el apartado anterior, para que un procesador realice un proceso se le debe suministrar en primer lugar un algoritmo adecuado. El procesador debe ser capaz de *interpretar* el algoritmo, lo que significa:

- comprender las instrucciones de cada paso,
- realizar las operaciones correspondientes.

Cuando el procesador es una computadora, el algoritmo se ha de expresar en un formato que se denomina *programa*, ya que el pseudocódigo o el diagrama de flujo no son comprensibles por la computadora, aunque pueda entenderlos cualquier programador. Un programa se escribe en un *lenguaje de programación* y las operaciones que conducen a expresar un algoritmo en forma de programa se llaman *programación*. Así pues, los lenguajes utilizados para escribir programas de computadoras son los *lenguajes de programación* y **programadores** son los escritores y diseñadores de programas. El proceso de traducir un algoritmo en *pseudocódigo* a un lenguaje de programación se denomina **codificación**, y el algoritmo escrito en un lenguaje de programación se denomina **código fuente**.

En la realidad la computadora no entiende directamente los lenguajes de programación sino que se requiere un programa que traduzca el código fuente a otro lenguaje que sí entiende la máquina directamente, pero muy complejo para las personas; este lenguaje se conoce como **lenguaje máquina** y el código correspondiente **código máquina**. Los programas que traducen el código fuente escrito en un lenguaje de programación —tal como C++— a código máquina se denominan **traductores**. El proceso de conversión de un algoritmo escrito en pseudocódigo hasta un programa ejecutable comprensible por la máquina, se muestra en la Figura 1.14.

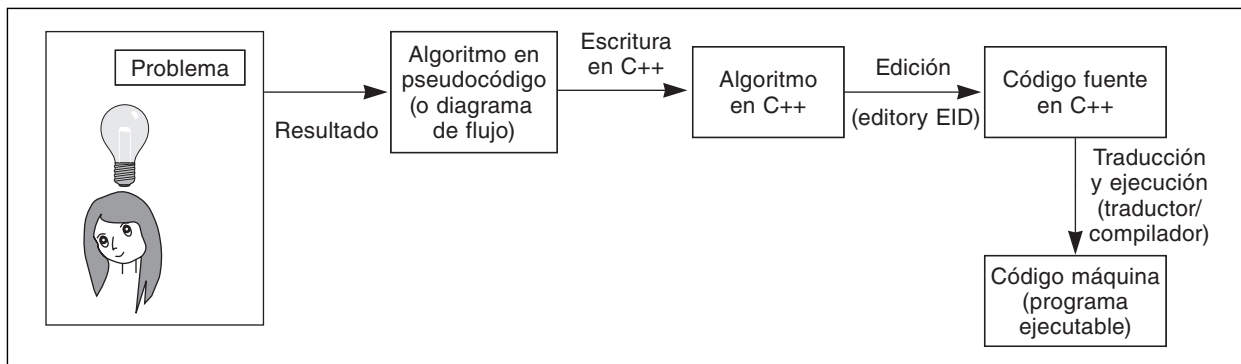


Figura 1.14. Proceso de transformación de un algoritmo en pseudocódigo en un programa ejecutable.

Hoy en día, la mayoría de los programadores emplean lenguajes de programación como C++, C, C#, Java, Visual Basic, XML, HTML, Perl, PHP, JavaScript..., aunque todavía se utilizan, sobre todo profesionalmente, los clásicos *COBOL*, *FORTRAN*, *Pascal* o el mítico BASIC. Estos lenguajes se denominan **lenguajes de alto nivel** y permiten a los profesionales resolver problemas convirtiendo sus algoritmos en programas escritos en alguno de estos lenguajes de programación.

Los **lenguajes de programación** se utilizan para escribir programas. Los programas de las computadoras modernas constan de secuencias de instrucciones que se codifican como secuencias de dígitos numéricos que podrán entender dichas computadoras. El sistema de codificación se conoce como **lenguaje máquina** que es el lenguaje nativo de una computadora. Desgraciadamente la escritura de programas en lenguaje máquina es una tarea tediosa y difícil ya que sus instrucciones son secuencias de 0 y 1 (*patrones de bit*, tales como 11110000, 01110011...) que son muy difíciles de recordar y manipular por las personas. En consecuencia, se necesitan lenguajes de programación “amigables con el programador” que permitan escribir los programas para poder “charlar” con facilidad

con las computadoras. Sin embargo, las computadoras sólo entienden las instrucciones en lenguaje máquina, por lo que será preciso traducir los programas resultantes a lenguajes de máquina antes de que puedan ser ejecutadas por ellas.

Cada lenguaje de programación tiene un conjunto o “juego” de instrucciones (acciones u operaciones que debe realizar la máquina) que la computadora podrá entender directamente en su código máquina o bien se traducirán a dicho código máquina. Las instrucciones básicas y comunes en casi todos los lenguajes de programación son:

- *Instrucciones de entrada/salida.* Instrucciones de transferencia de información entre dispositivos periféricos y la memoria central, tales como "leer de..." o bien "escribir en...".
- *Instrucciones de cálculo.* Instrucciones para que la computadora pueda realizar operaciones aritméticas.
- *Instrucciones de control.* Instrucciones que modifican la secuencia de la ejecución del programa.

Además de estas instrucciones y dependiendo del procesador y del lenguaje de programación existirán otras que conformarán el conjunto de instrucciones y junto con las reglas de sintaxis permitirán escribir los programas de las computadoras. Los principales tipos de lenguajes de programación son:

- *Lenguajes máquina.*
- *Lenguajes de bajo nivel (ensambladores).*
- *Lenguajes de alto nivel.*



Figura 1.15. Diferentes sistemas operativos: Windows Vista (izquierda) y Red Hat Enterprise Linux 4.

1.9.1. Traductores de lenguaje: el proceso de traducción de un programa

El proceso de traducción de un programa fuente escrito en un lenguaje de alto nivel a un lenguaje máquina comprensible por la computadora, se realiza mediante programas llamados “traductores”. Los **traductores de lenguaje** son programas que traducen a su vez los programas fuente escritos en lenguajes de alto nivel a código máquina. Los traductores se dividen en **compiladores** e **intérpretes**.

Intérpretes

Un *intérprete* es un traductor que toma un programa fuente, lo traduce y, a continuación, lo ejecuta. Los programas intérpretes clásicos como BASIC, prácticamente ya no se utilizan, más que en circunstancias especiales. Sin embargo, está muy extendida la versión interpretada del lenguaje Smalltalk, un lenguaje orientado a objetos puro. El sistema de traducción consiste en: traducir la primera sentencia del programa a lenguaje máquina, se detiene la traducción, se ejecuta la sentencia; a continuación, se traduce la siguiente sentencia, se detiene la traducción, se ejecuta la sentencia y así sucesivamente hasta terminar el programa (Figura 1.16).

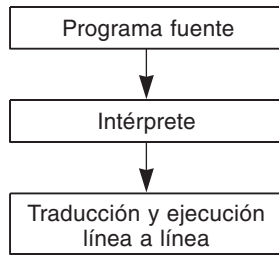


Figura 1.16. Intérprete.

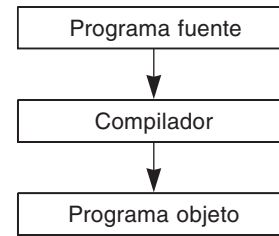


Figura 1.17. La compilación de programas.

Compiladores

Un *compilador* es un programa que traduce los programas fuente escritos en lenguaje de alto nivel a lenguaje máquina. La traducción del programa completo se realiza en una sola operación denominada **compilación** del programa; es decir, se traducen todas las instrucciones del programa en un solo bloque. El programa compilado y depurado (eliminados los errores del código fuente) se denomina *programa ejecutable* porque ya se puede ejecutar directamente y cuantas veces se desee; sólo deberá volver a compilarse de nuevo en el caso de que se modifique alguna instrucción del programa. De este modo el programa ejecutable no necesita del compilador para su ejecución. Los traductores de lenguajes típicos más utilizados son: **C**, **C++**, **Java**, **C#**, **Pascal**, **FORTRAN** y **COBOL** (Figura 1.17).

1.9.2. La compilación y sus fases

La *compilación* es el proceso de traducción de programas fuente a programas objeto. El programa objeto obtenido de la compilación ha sido traducido normalmente a código máquina.

Para conseguir el programa máquina real se debe utilizar un programa llamado *montador* o *enlazador* (*linker*). El proceso de montaje conduce a un programa en lenguaje máquina directamente ejecutable (Figura 1.18).

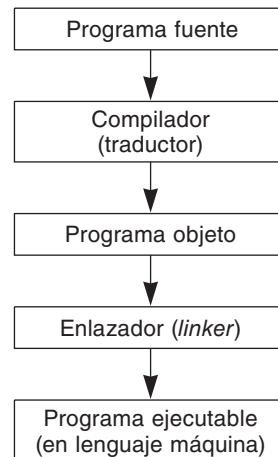


Figura 1.18. Fases de la compilación.

El proceso de ejecución de un programa escrito en un lenguaje de programación y mediante un compilador suele tener los siguientes pasos:

1. Escritura del *programa fuente* con un *editor* (programa que permite a una computadora actuar de modo similar a una máquina de escribir electrónica) y guardarlo en un dispositivo de almacenamiento (por ejemplo, un disco).
2. Introducir el programa fuente en memoria.

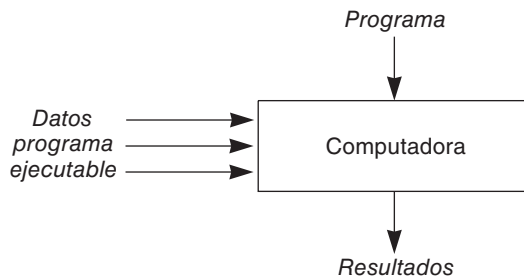


Figura 1.19. Ejecución de un programa.

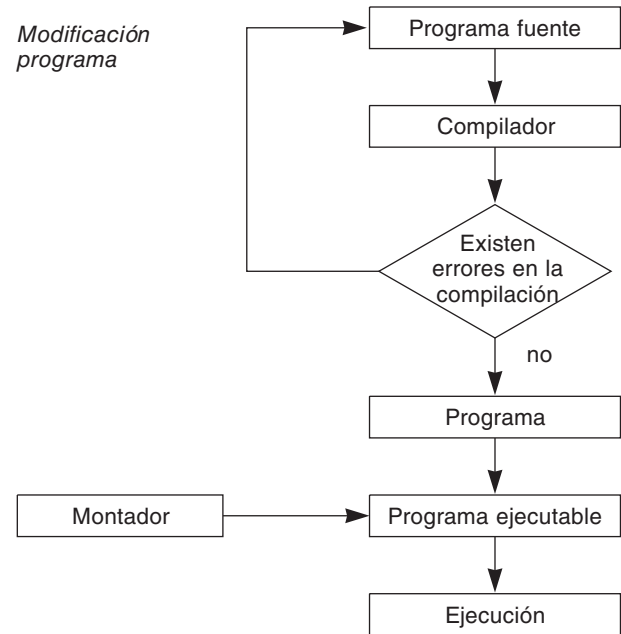


Figura 1.20. Fases de ejecución de un programa.

3. *Compilar* el programa con el compilador seleccionado.
4. *Verificar y corregir errores de compilación* (listado de errores).
5. Obtención del programa *objeto*.
6. El enlazador (*linker*) obtiene el *programa ejecutable*.
7. Se ejecuta el programa y, si no existen errores, se tendrá la salida del programa.

El proceso de ejecución se muestra en las Figuras 1.19 y 1.20.

1.9.3. Evolución de los lenguajes de programación

En la década de los cuarenta cuando nacían las primeras computadoras digitales el lenguaje que se utilizaba para programar era el *lenguaje máquina* que traducía directamente el código máquina (código binario) comprensible para las computadoras. Las instrucciones en lenguaje máquina dependían de cada computadora y debido a la dificultad de su escritura, los investigadores de la época simplificaron el proceso de programación desarrollando sistemas de notación en los cuales las instrucciones se representaban en formatos *nemónicos* (nemotécnicos) en vez de en formatos numéricos que eran más difíciles de recordar. Por ejemplo, mientras la instrucción

```
Mover el contenido del registro 4 al registro 8
```

se podía expresar en lenguaje máquina como

```
4048      o bien    0010  0000 0010 1000
```

en código *nemotécnico* podía aparecer como

```
MOV R5, R6
```

Para convertir los programas escritos en código nemotécnico a lenguaje máquina, se desarrollaron programas ensambladores (*assemblers*). Es decir, los **ensambladores** son programas que traducen otros programas escritos en código nemotécnico en instrucciones numéricas en lenguaje máquina que son compatibles y legibles por la máquina. Estos programas de traducción se llaman ensambladores porque su tarea es ensamblar las instrucciones reales de la

máquina con los nemotécnicos e identificadores que representan las instrucciones escritas en ensamblador. A estos lenguajes se les denominó de segunda generación, reservando el nombre de primera generación para los lenguajes de máquina.

En la década de los cincuenta y sesenta comenzaron a desarrollarse lenguajes de programación de tercera generación que diferían de las generaciones anteriores en que sus instrucciones o primitivas eran de alto nivel (comprensibles por el programador, como si fueran lenguajes naturales) e **independientes de la máquina**. Estos lenguajes se llamaron **lenguajes de alto nivel**. Los ejemplos más conocidos son **FORTRAN** (FORmula TRANslator) que fue desarrollado para aplicaciones científicas y de ingeniería, y **COBOL** (COMmon Business-Oriented Language), que fue desarrollado por la U.S. Navy de Estados Unidos, para aplicaciones de gestión o administración. Con el paso de los años aparecieron nuevos lenguajes tales como **Pascal**, **BASIC**, **C**, **C++**, **Ada**, **Java**, **C#**, **HTML**, **XML**...

Los lenguajes de programación de alto nivel se componen de un conjunto de instrucciones o primitivas más fáciles de escribir y recordar su función que los lenguajes máquina y ensamblador. Sin embargo, los programas escritos en un lenguaje de alto nivel, como C o Java necesitan ser traducidos a código máquina; para ello se requiere un programa denominado **traductor**. Estos programas de traducción se denominaron técnicamente, **compiladores**. De este modo existen compiladores de C, FORTRAN, Pascal, Java, etc.

También surgió una alternativa a los traductores compiladores como medio de implementación de lenguajes de tercera generación que se denominaron **intérpretes**²³. Estos programas eran similares a los traductores excepto que ellos ejecutaban las instrucciones a medida que se traducían, en lugar de guardar la versión completa traducida para su uso posterior. Es decir, en vez de producir una copia de un programa en lenguaje máquina que se ejecuta más tarde (este es el caso de la mayoría de los lenguajes, C, C++, Pascal, Java...), un intérprete ejecuta realmente un programa desde su formato de alto nivel, instrucción a instrucción. Cada tipo de traductor tiene sus ventajas e inconvenientes, aunque hoy día prácticamente los traductores utilizados son casi todos compiladores por su mayor eficiencia y rendimiento.

Sin embargo, en el aprendizaje de programación se suele comenzar también con el uso de los lenguajes algorítmicos, similares a los lenguajes naturales, mediante instrucciones escritas en *pseudocódigo* (o *seudocódigo*) que son palabras o abreviaturas de palabras escritas en inglés, español, portugués, etc. Posteriormente se realiza la conversión al lenguaje de alto nivel que se vaya a utilizar realmente en la computadora, tal como C, C++ o Java. Esta técnica facilita la escritura de algoritmos como paso previo a la programación.

1.9.4. Paradigmas de programación

La evolución de los lenguajes de programación ha ido paralela a la idea de paradigma de programación: enfoques alternativos a los procesos de programación. En realidad un **paradigma de programación** representa fundamentalmente enfoques diferentes para la construcción de soluciones a problemas y por consiguiente afectan al proceso completo de desarrollo de software. Los paradigmas de programación clásicos son: *procedimental* (o *imperativo*), *funcional*, *declarativo* y *orientado a objetos*. En la Figura 1.21 se muestra la evolución de los paradigmas de programación y los lenguajes asociados a cada paradigma [BROOKSHEAR 04]²⁴.

Lenguajes imperativos (procedimentales)

El **paradigma imperativo o procedimental** representa el enfoque o método tradicional de programación. Un lenguaje imperativo es un conjunto de instrucciones que se ejecutan una por una, de principio a fin, de modo secuencial excepto cuando intervienen instrucciones de salto de secuencia o control. Este paradigma define el proceso de programación como el desarrollo de una secuencia de órdenes (*comandos*) que manipulan los datos para producir los resultados deseados. Por consiguiente, el paradigma imperativo señala un enfoque del proceso de programación mediante la realización de un algoritmo que resuelve de modo manual el problema y a continuación expresa ese algoritmo como una secuencia de órdenes. En un lenguaje procedimental cada instrucción es una orden u órdenes para que la computadora realice alguna tarea específica.

²³ Uno de los intérpretes más populares en las décadas de los setenta y ochenta, fue **BASIC**.

²⁴ J. Glenn Brookshear, *Computer Science: An overview, Eighth edition*, Boston (EE.UU.): Pearson/Addison Wesley, 2005, p. 230. Obra clásica y excelente para la introducción a la informática y a las ciencias de la computación en todos sus campos fundamentales. Esta obra se recomienda a todos los lectores que deseen profundizar en los diferentes temas tratados en este capítulo y ayudará considerablemente al lector como libro de consulta en su aprendizaje en programación.

Los lenguajes de programación procedimentales, por excelencia, son **FORTRAN**, **COBOL**, **Pascal**, **BASIC**, **ALGOL**, **C** y **Ada** (aunque sus últimas versiones ya tienen un carácter completamente orientado a objetos).

Lenguajes declarativos

En contraste con el paradigma imperativo el **paradigma declarativo** solicita al programador que describa el problema en lugar de encontrar una solución algorítmica al problema; es decir, un lenguaje declarativo utiliza el principio del razonamiento lógico para responder a las preguntas o cuestiones consultadas. Se basa en la *lógica formal* y en el *cálculo de predicados de primer orden*. El razonamiento lógico se basa en la deducción. El lenguaje declarativo por excelencia es **Prolog**.

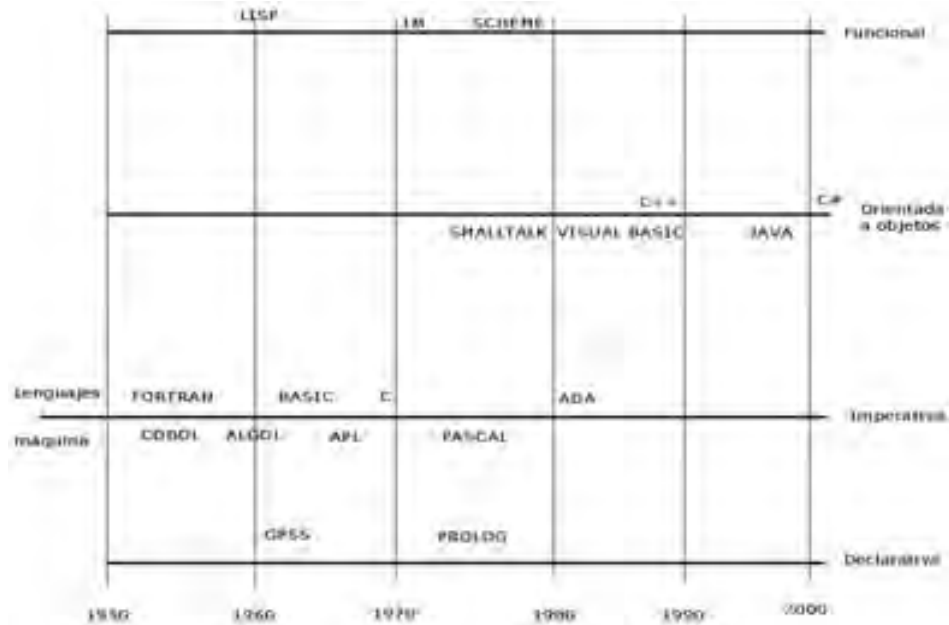


Figura 1.21. Paradigmas de programación (evolución de lenguajes).

Lenguajes orientados a objetos

El paradigma orientado a objetos se asocia con el proceso de programación llamado **programación orientada a objetos (POO)**²⁵ consistente en un enfoque totalmente distinto al proceso procedimental. El enfoque orientado a objetos guarda analogía con la vida real. El desarrollo de software OO se basa en el diseño y construcción de objetos que se componen a su vez de datos y operaciones que manipulan esos datos. El programador define en primer lugar los objetos del problema y a continuación los datos y operaciones que actuarán sobre esos datos. Las ventajas de la programación orientada a objetos se derivan esencialmente de la estructura modular existente en la vida real y el modo de respuesta de estos módulos u objetos a mensajes o eventos que se producen en cualquier instante.

Los orígenes de la POO se remontan a los Tipos Abstractos de Datos como parte constitutiva de una estructura de datos. En este libro se dedicará un capítulo completo al estudio del **TAD** como origen del concepto de programación denominado **objeto**.

C++ lenguaje orientado a objetos, por excelencia, es una extensión del lenguaje C y contiene las tres propiedades más importantes: *encapsulamiento*, *herencia* y *polimorfismo*. Smalltalk es otro lenguaje orientado a objetos muy potente y de gran impacto en el desarrollo del software orientado a objetos que se ha realizado en las últimas décadas.

Hoy día **Java** y **C#** son herederos directos de C++ y C, y constituyen los lenguajes orientados a objetos más utilizados en la industria del software del siglo XXI. **Visual Basic** y **VB.Net** son otros lenguajes orientados a objetos, no tan potentes como los anteriores pero extremadamente sencillos y fáciles de aprender.

²⁵ Si desea profundizar en este tipo de programación existen numerosos y excelentes libros que puede consultar en la Bibliografía.

1.10. BREVE HISTORIA DE LOS LENGUAJES DE PROGRAMACIÓN

La historia de la computación ha estado asociada indisolublemente a la aparición y a la historia de lenguajes de programación de computadoras²⁶. La Biblia de los lenguajes ha sido una constante en el desarrollo de la industria del software y en los avances científicos y tecnológicos. Desde el año 1642 en que Blaise Pascal, inventó *La Pascalina*, una máquina que ayudaba a contar mediante unos dispositivos de ruedas, se han sucedido numerosos inventos que han ido evolucionando, a medida que se programaban mediante códigos de máquina, lenguajes ensambladores, hasta llegar a los lenguajes de programación de alto nivel en los que ya no se dependía del hardware de la máquina sino de la capacidad de abstracción del programador y de la sintaxis, semántica y potencia del lenguaje.

En la década de los cincuenta, IBM diseñó el primer lenguaje de programación comercial de alto nivel y concebido para resolver problemas científicos y de ingeniería (FORTRAN, 1954). Todavía hoy, muchos científicos e ingenieros siguen utilizando FORTRAN en sus versiones más recientes FORTRAN 77 y FORTRAN 90. En 1959, la doctora y almirante, Grace Hopper, lideró el equipo que desarrolló COBOL, el lenguaje por excelencia del mundo de la gestión y de los negocios hasta hace muy poco tiempo; aunque todavía el mercado sigue demandando programadores de COBOL ya que numerosas aplicaciones comerciales siguen corriendo en este lenguaje.

Una enumeración rápida de lenguajes de programación que han sido o son populares y los años en que aparecieron es la siguiente:

Década 50	Década 60	Década 70	Década 80	Década 90	Década 00
FORTRAN (1954)	BASIC (1964)	Pascal (1970)	C++ (1983)	Java (1997)	C# (2000)
ALGOL 58 (1958)	LOGO (1968)	C (1971)	Eiffel (1986)		
LISP (1958)	Simula 67 (1967)	Modula 2 (1975)	Perl (1987)		
COBOL (1959)	Smalltalk (1969)	Ada (1979)			

Programación de la Web

Si después o en paralelo de su proceso de aprendizaje en fundamentos y metodología de la programación desea practicar no sólo con un lenguaje tradicional como Pascal, C, C++, Java o C#, sino introducirse en lenguajes de programación para la Web, enumeramos a continuación los más empleados en este campo.

Los programadores pueden utilizar una amplia variedad de lenguajes de programación, incluyendo C y C++ para escribir aplicaciones Web. Sin embargo, algunas herramientas de programación son, particularmente, útiles para desarrollar aplicaciones Web:

- **HTML**, técnicamente es un lenguaje de descripción de páginas más que un lenguaje de programación. Es el elemento clave para la programación en la Web.
- **JavaScript**, es un lenguaje interpretado de guionado (scripting) que facilita a los diseñadores de páginas Web añadir guiones a páginas Web y modos para enlazar **esas páginas**.
- **VBScript**, la respuesta de Microsoft a JavaScript basada en VisualBasic.
- **Java**, lenguaje de programación, por excelencia, de la Web.
- **ActiveX**, lenguaje de Microsoft para simular a algunas de las características de Java.
- **C#**, el verdadero competidor de Java y creado por Microsoft.
- **Perl**, lenguaje interpretado de guionado (scripting) idóneo para escritura de texto.
- **XML**, lenguaje de marcación que resuelve todas las limitaciones de HTML y ha sido el creador de una nueva forma de programar la Web. Es el otro gran lenguaje de la Web.
- **AJAX**, es el futuro de la Web. Es una mezcla de JavaScript y XML. Es la espina dorsal de la nueva generación Web 2.0.

²⁶ Si desea una breve historia pero más detallada de los lenguajes de programación más utilizados por los programadores profesionales tanto para aprendizaje como para el desarrollo profesional puede consultarlo en la página web del libro: www.mhe.es/joyanes.

En el sitio Web de la editorial O'Reilly puede descargarse un póster (en PDF) con una magnífica y fiable Historia de los Lenguajes de Programación: www.oreilly.com/news/graphics/prog_lang_poster.pdf

RESUMEN

Una computadora es una máquina para procesar información y obtener resultados en función de unos datos de entrada.

Hardware: parte física de una computadora (dispositivos electrónicos).

Software: parte lógica de una computadora (programas).

Las computadoras se componen de:

- Dispositivos de Entrada/Salida (E/S).
- Unidad Central de Proceso (Unidad de Control y Unidad Lógica y Aritmética).
- Memoria central.
- Dispositivos de almacenamiento masivo de información (memoria auxiliar o externa).

El *software del sistema* comprende, entre otros, el sistema operativo **Windows**, **Linux**, en computadoras personales y los lenguajes de programación. Los *lenguajes de programación* de alto nivel están diseñados para hacer más fácil la

escritura de programas que los lenguajes de bajo nivel. Existen numerosos lenguajes de programación cada uno de los cuales tiene sus propias características y funcionalidades, y normalmente son más fáciles de transportar a máquinas diferentes que los escritos en lenguajes de bajo nivel.

Los programas escritos en lenguaje de alto nivel deben ser traducidos por un compilador antes de que se puedan ejecutar en una máquina específica. En la mayoría de los lenguajes de programación se requiere un compilador para cada máquina en la que se desea ejecutar programas escritos en un lenguaje específico...

Los lenguajes de programación se clasifican en:

- *Alto nivel:* Pascal, FORTRAN, Visual Basic, C, Ada, Modula-2, C++, Java, Delphi, C#, etc.
- *Bajo nivel:* Ensamblador.
- *Máquina:* Código máquina.
- *Diseño de Web:* SMGL, HTML, XML, PHP...

Los programas traductores de lenguajes son:

- *Compiladores.*
- *Intérpretes.*