Understanding Your Computer's Personality

Can computers have personalities? Not in the human sense. But most of us have heard of Macs and PCs, and if you've used both types of desktop computer, you know that they "behave" differently, in subtle and not-so-subtle ways. This can be considered a personality of sorts, and the source of this personality is something that computer specialists call an *operating system.* In this essay, we'll explain what an operating system is, and how it determines your computer's personality.

Basic Notions

Computers are electronic devices that can carry out simple manipulations of information according to a prepared set of instructions. The information, stored inside the computer's circuits by turning individual circuits on and off, is collectively referred to as *data.* The set of instructions that directs the computer's work is also stored inside the computer, and is called a *program*. *Hardware* is the term used to describe the actual electronic circuits, fans, wires, and so on that make up the physical computer. *Software* is the term used to collectively refer to the various programs one can use to tell the computer what to do. Software is prepared in advance by human beings called programmers. Data is introduced, or *input,* into the computer by human beings called *users,* via a keyboard or mouse or other device. The computer *outputs* the results of its work on a screen, on paper, or in various other ways. When a computer is following the instructions of a program, it is said to be *running* or *executing* the program.

A key distinction between hardware and software is that hardware is a set of physical components that cannot easily be changed without taking the computer apart, whereas software is just sets of instructions that can be loaded into the computer on an as-needed basis very rapidly. Thus, anything built into the hardware cannot change, whereas anything controlled by the software can be changed by simply modifying the software (which can be done in a few seconds).

Virtually all desktop computers are built from the same interchangeable pieces of hardware. The only thing that really distinguishes one desktop computer from another is the type of software that it runs. An Apple Macintosh desktop computer behaves somewhat differently from a Dell desktop computer because it runs different software.

The data and software used by a computer must be stored inside its electronic memory before it can use them. This memory, however, is erased whenever the power is turned off. Computers therefore have provisions for automatically loading software into the computer when they are turned on, and they are also equipped with storage devices, called *disk drives,* that are capable of storing information even when the computer is turned off. The computer is designed to load software from the disk drive as soon as it is turned on, and additional software can be loaded as needed from the disk as the computer operates. This same disk also stores the data used by the computer whenever the computer is not actually working on it.

The Importance of Software

Computers are intended to be extremely general-purpose devices, capable of doing just about any work involving any type of data. To this end, the computer hardware is designed to carry out only extremely simple operations (add, subtract, move, input, output, and so on) in response to equally simple instructions contained in a program. The individual operations are simple and primitive, but since a typical computer today can execute several

billion instructions each second, it is possible to do just about anything with a sufficiently complex set of instructions. Conversely, without instructions, a computer does absolutely nothing at all—it becomes an expensive paperweight. This means that a computer is completely useless without software, but it also means that the computer offers tremendous flexibility, since so much of what it does is controlled by software that can be easily modified.

Since the computer is helpless without instructions, it is necessary for it to have instructions to execute at all times. Worse yet, even the simplest functions of a computer require a great many instructions to carry out—accepting a letter typed typed by a human being on the keyboard of the computer, for example, involves the execution of tens of thousands of individual instructions. It follows that the computer must have various programs constantly available that tell it how to accept input from the keyboard or mouse, how to display output on the computer screen, how to communicate with the Internet, if the computer is connected to the Net, and so on.

Getting the computer to correctly handle all the devices with which it is equipped (mouse, keyboard, screen, etc.), requires many millions of instructions. These instructions are pretty much the same no matter what the computer is actually doing. Preparing the instructions is an extremely time-consuming task for human programmers.

The Birth of Operating Systems

In the olden days, every program run on a computer, no matter what its purpose—accounting, weather prediction, aircraft simulation, whatever—included thousands or millions of instructions that told the computer how to carry out basic functions like input and output, as we've described above. When a given program was needed, the computer's memory was cleared, then the necessary program was loaded into its memory and the computer was started. Each program contained all the instructions it needed, not only for its nominal purpose (such as processing the payroll or generating invoices) but also for all the mundane functions it used, such as printing information on a printer or showing it on a screen.

The first computers were very slow, and so the fact that they had to be stopped, cleared, and reloaded for each program was not a big issue. Only one person could use the computer at a time. He would would put the program he needed and the data he wanted to process onto punched cards or reels of magnetic tape, and would then carry them into the computer room, and he would load them into the computer and press the START button to get the computer going. When the computer finished, it would stop, and the human being would remove the results from the computer, clear its memory, and it would be ready for its next job.

This was all well and good as long as computers were very slow, but they increased in speed rapidly, and their cost rose correspondingly. Eventually the computer became so fast that it would execute programs in a few seconds, and then would sit idle for minutes or hours while a human being prepared the next program for use. Computers that cost millions of dollars were 99% idle because they spent most of their time waiting for a human being to load or unload a program. Something had to be done.

Eventually it occurred to computer vendors and users that a special program could be prepared that would do nothing but load and execute other programs. This special program would be loaded once into the computer when it was turned on, and would then run continuously thereafter; it would load other programs into the computer, load the input they required, execute the programs, collect the results, and then output the results. It would then repeat the process for the next program. Since the computer was already much faster than its human masters, using a special program to run all other programs vastly increased the efficiency of the computer and made it possible to keep the computer almost continuously busy with useful work.

This same special program also provided all the software necessary to perform basic functions such as reading a magnetic tape or printing something on a printer. The other programs that were loaded by this special program no longer needed to contain instructions for these basic functions them-

selves; all they had to do was communicate their desires to the special program, which would then execute the necessary instructions to read a tape block or print a page.

The special program that ran continuously in the computer handled most of the tasks formerly handled by human operators. As a result, the special program came to be called the *operating system* (or *OS),* and that term is still in use today, decades later.

What Operating Systems Do

All modern computers have operating systems. Operating systems are programs that run continuously in the computer. They fulfill several functions. They organize the operation of the computer so that it remains consistent no matter how the computer is being used. They provide useful services to other programs that are run on the computer. They ensure that the security of the computer system is maintained. They schedule the execution of other programs in such a way that computer time is not wasted. And they organize the storage of information on the computer's *disk drives,* the devices used to store large amounts of information over the long term.

The programs contained within the operating system are collectively referred to as *system software.* All other programs—meaning programs for word processing, games, business programs, and so on—are collectively referred to as *application software.*

The operating system is always present in the computer. It stays in the background, constantly managing the computer hardware, organizing and monitoring the execution of application programs, and generally carrying out a vast number of tedious but necessary functions. The operating system itself doesn't perform any useful functions from a human standpoint—it doesn't balance checkbooks, check e-mail, or surf the Web—but it provides essential services to the application programs that *do* perform these useful functions. In modern computers, operating systems are mandatory, and all application software is written with the assumption that it will run under the control of an operating system.

The Personality of an OS

The constant presence of an operating system in a computer and the influence it has on application software create a strong consistency in the behavior of the computer. This consistency represents the "look and feel" or "personality" of the computer. While the operating system itself doesn't usually interact with the human user of a computer in an obvious way, it constantly influences the way that other application programs interact, and so it creates an "atmosphere" that is highly specific to that operating system.

The hardware in your desktop PC is capable of running many different operating systems. If you change the operating system in the computer, the personality of the computer may dramatically change as well. The hardware is the same, but the software is different, and since software has almost unlimited influence over what a computer actually does, changing the operating system—the software that constantly runs in the background on the computer—also changes the look and feel of your interaction with the computer.

For example, application programs create windows with which you can interact under a number of operating systems. The Microsoft Windows operating system, which runs on 95% of desktop PCs, has a consistent way of drawing windows on the screen of the computer, and no matter what application you run under this operating system, the style of the windows tends to look the same. The Mac OS X operating system, which can run on the same hardware as Microsoft Windows, also creates windows for application programs—but the style of the windows created by OS X is noticeably different from that of windows created by Microsoft Windows. The *look and feel* of the operating system is significantly different, even though the overall way in which the computer is used remains the same, and even though the application program itself may be similar.

Microsoft Word is a popular application program that is available for both Microsoft Windows and Mac OS X. The program works almost identically under both operating systems, but it looks somewhat different. The difference in the look of the

program is largely due to the difference in the operating systems under which it runs.

In some cases, changing the operating system running on a computer can have a much more dramatic effect on the way it behaves. Operating systems like OS X and Windows share the concept of a *graphic user interface,* or GUI, that includes windows drawn on the screen, icons, a mouse cursor, and many other features. But no law requires that an operating system work in this way, and some operating systems work in a very different way. The UNIX operating system can run on the same hardware as Windows or OS X, but the basic version of UNIX doesn't have a GUI at all; it just displays plain text on a screen.

Applications and Operating Systems

Today, all application programs are designed to run under the control of an operating system. And since no two operating systems are identical, typically a separate version of an application must be written for each operating system under which it will run. Adobe Photoshop is a popular application program that runs under both Windows and Mac OS X, but the program actually exists in two separate versions, one compatible with the Windows operating system, and the other compatible with the Mac operating system. The two versions behave in much the same way from the user standpoint, but they are not interchangeable.

This situation exists because the interface between the operating system and application programs is different for every OS, and all application programs communicate extensively with the operating system while they are running. A call to the operating system to put a window on the screen isn't exactly the same in Windows as it is in OS X, so separate versions of the application must be prepared that use the correct call for the correct OS.

A very important consequence of this is that the type of operating system installed on a computer controls the applications that can run on that computer. For some operating systems, such as Windows or OS X, there are tens of thousands or even hundreds of thousands of applications available. For other, more obscure operating systems, there may be hardly any applications available. Often the choice of operating system for a computer is determined by the applications one wishes to run. If you need to use application X, you need an operating system for which application X is available. One reason why Microsoft Windows is used on the vast majority of desktop computers is that there are more applications available for that operating system than for any other operating system in the world.

Operating Systems and Security

One of the many services provided by an operating system is security. Indeed, the only reason modern computers provide any security at all is that modern operating systems have security features built in.

When a computer runs with an operating system, the operating system is normally *privileged,* meaning that it can execute certain instructions that other programs cannot. This is enforced by the computer hardware. The hardware runs in two modes, *privileged* or *supervisor* mode, and *non-privileged* or *user* mode. When the computer is first turned on, it runs in privileged mode as it loads the operating system. Thereafter, the operating system switches the mode of the hardware to non-privileged whenever it passes control to an application program. Thus, the operating system always executes in a privileged mode, whereas application programs always execute in a non-privileged mode. The instructions that are available only in privileged mode include all the instructions that move data into and out of the computer, as well as instructions that affect the overall stability of the system. Only the operating system can use these special instructions.

Because of these combined hardware and software controls, ordinary application programs cannot damage a computer, if the operating system is designed to keep the computer safe. If an application program attempts to access something without proper permission from the operating system, the OS will intervene, by *aborting* (stopping) the application program. While application programs are free to do whatever they wish in their own little world (created for them by the OS), they cannot

touch other parts of the system outside this world. Applications cannot touch each other except under strictly controlled conditions, and always through the intermediary of the operating system, never directly. Likewise, they cannot touch the computer hardware directly; if an application wants to use some part of the computer hardware (the display, a printer, etc.), it must do so by making a request to the operating system, which then validates the request and carries it out, if it is permitted.

User Contexts

The operating system also enforces the concept of *user contexts.* User contexts are sets of privileges assigned to application programs based in part on the identity of the human being using the computer. The human user is given a *user account* that provides the user context in which her application programs will run.

The operating system identifies a human user by asking her to identify herself with a user or account name and a password, or by some other means (smart card, fingerprint, etc.). Once the OS has determined that the user is authorized to use the computer, it prepares a user context for her. Any programs executed by the user on the computer are executed in her user context. The privileges enjoyed by the programs are limited to those for which the human user is authorized. A program that attempts to use privileges for which it has no authority is aborted.

All operating systems that recognize user contexts also include the notion of an all-powerful *superuser*—also variously called the *administrator, root, master user, sysadmin,* etc., depending on the terminology of the specific OS. This special user has all privileges, and can run any program or carry out any operation on the computer. The operating system never refuses anything to the superuser. The superuser exists because there has to be some way to perform maintenance operations on the computer, and only a user with all privileges can carry out all the operations that might be needed.

Ordinary users of a computer normally use ordinary user accounts when accessing it. A user who is also the owner or administrator of a computer may have two accounts, one being the superuser account, the other being an ordinary user account; and normally she will use only the ordinary user account, saving the superuser account for occasions that really require special privileges.

Operating-system security is important because it helps protect a computer against untrustworthy or defective application software. Since no application program can obtain the type of privileges that would be necessary to do serious damage to the computer system, it is not necessary to completely trust such a program before using it. The operating system is trustworthy by definition, and it controls all access to sensitive functions. In a computer without an operating system, application programs can do anything they want once they begin to run.

Organizing Files

Another advantage provided by most operating systems is file organization, *i.e.,* a structured way of organization information permanently stored on the computer's disk drives.

Computer disk drives are simply storage devices that can hold hundreds of billions of bytes of information (the equivalent of a thousand encyclopedias or more). From a hardware standpoint, the space on a disk is just one vast expanse, without any particular organization; but to make productive use of so much space, some sort of organization is a good idea—and an operating system provides this.

In a computer without an operating system, every application program has complete control of the computer—and complete access to all the space on the disk drives. Since each program is likely to have its own idea of how to use the disk drives, there will almost certainly be a conflict between the way different application programs organize data on the disks; and the information recorded on the disks by one program is likely to be unrecognizable to all other programs. Different programs may even destroy each other's data as they run, since they probably organize information on disk in conflicting ways. The solution to all of these problems is an operating system.

When an operating system is present, only the operating system has direct access to the comput-

er's disk drives. All application programs must request access to disks via the operating system. The operating system organizes information on the disks in a coherent way and imposes this organization on all application programs.

Standard Desktop File Organization

In most operating systems, including those for desktop computers, the information on disk is organized into *files* and *folders*. A file is just an area on the disk set aside for a specific purpose, such as storage of a document, a picture, a spreadsheet, or whatever. A folder is a collection of files. Folders can themselves contain other folders, making it possible to create a tree-like structure of folders and subfolders. Each folder at any level can also contain files. Both files and folders are given names, so that they can be easily identified.

Application programs may read, write, create, modify, and delete (erase) files and folders on the disks, by making appropriate requests to the operating system. No application is ever allowed to access the disk directly. Instead, the desired operation is requested of the operating system, which then validates the request (to make sure that it doesn't violate any security precautions), and then carries it out, if the requesting application is authorized to ask for the operation.

By forcing all disk operations through the operating system, the integrity of information on the disk can be assured. Operating systems often have extensive security controls in place for access to disks. For example, folders and files are associated with specific owners, meaning that a given file or folder "belongs" to a particular user account. Users associated with this account may freely access the file or folder, but other users may or may not be allowed access, depending on what the owner of the file or folder decides. The various authorizations that are possible are quite complex, making it easy to carefully control the way in which files and folders may be used by individual applications.

For instance, a given application running under a particular user account might be allowed to read the contents of a file belonging to another user account, but it might not be allowed to write or delete that file.

System files are files owned and manipulated by the operating system itself. Nobody is allowed to access these files except the operating system (or the superuser). These files contain information critical to the stability and security of the entire computer, and so they are hidden from all ordinary users, and any attempt by an ordinary user program to access them causes the program to be aborted.

Because the organization of information on disks is specific to each operating system, it isn't possible to use one operating system to access information on disks organized by another operating system (unless special utility programs are provided to do this). So if you install Microsoft Windows on your desktop computer, you cannot subsequently access the information you've saved in files and folders under Windows with any other operating system. Indeed, since the operating system itself is stored on the disk, once you've installed it, you generally cannot use any other operating system without erasing the first one from the disk and installing a new one.

Running as Root

Most people using desktop computers "run as root," meaning that they use the computer's system administrator account for all their work. This is very convenient, but it is also dangerous. When everything on the computer is done using the "root" or superuser account, there are no restrictions on access to the computer—and if a program executed under the superuser account happens to attempt something malicious, nothing will stop it.

Very paranoid users of desktop computers will often create an ordinary user account for their regular use of the computer. This adds a lot of protection for the computer, but it also makes use of the computer more awkward. When you use a single account (especially the superuser account) for all work on a computer, the security system of the computer is effectively hidden, since all files and folders are owned by the same account, and all access to the computer is unrestricted (if you run as the superuser). If you use a separate account for ordinary day-to-day work, files and folders may or may not belong to your account, and you may or may not be allowed to do certain things. This pre-

vents untrustworthy programs from damaging your computer, but it also may prevent you from doing useful work in some cases. This is particularly true for desktop computers because many common application programs were written before desktop operating systems included security features, and they assume that they have unrestricted access to everything—if you run such programs under a non-privileged account, they may crash as soon as they try to do something that requires special privileges.

Desktop Operating Systems

As we have previously pointed out, there are many operating systems available for desktop computers. A few of the most commonly used desktop operating systems are briefly described below.

Microsoft Windows

The most popular desktop operating system today is called Windows, and is a product of Microsoft Corporation. About nineteen out of twenty desktop computers today run this operating system, which is often installed directly at the factor while the computer is being built.

Windows actually exists in several versions, which are moderately backward-compatible with each other. (*Backward compatibility* means that you can replace an early version with a later version without losing any information, but you cannot replace a later version with an early version.)

The earliest versions of Windows, first released in the middle of the 1980s, were very primitive and are obsolete today. A later group of Windows operating systems, called the Windows 9x family and including Windows 95, Windows 98, Windows 98SE, and Windows ME, was released beginning in 1995 and is now obsolete as well, although it is still in widespread use around the world. The current Windows family, which began with Windows NT in 1994 and continues with Windows XP and Windows 2003 today, represents dramatic advances over its predecessors, although the basic look and feel of the operating system has remained the same since the very earliest versions, and is backward-compatible with them for the most part.

A key advantage of Windows is that more application programs are available for Windows than for any other desktop operating system—more than a quarter-million today. Another advantage is its familiarity: most people who have used desktop computers have used them with the Windows operating system, and so it is usually not necessary to explicitly train people in the use of Windows.

A disadvantage of Windows is its high cost, although the cost is not that unreasonable for a product of such complexity, and often the cost is included with the price of a new computer (since Windows is frequently preinstalled at the factory). Another disadvantage is the huge complexity of the operating system, which grows dramatically with each new version, requiring constant computer upgrades just to keep it running at an acceptable speed. Finally, because Windows is so widely used, it is a preferred target of cybercriminals, who constantly scheme to infect Windows systems with viruses and worms, and continually attack Windows systems over the Internet.

Nevertheless, Windows is probably the best choice for desktop computer users who are not computer specialists, with the advantages outweighing the disadvantages.

The Mac Operating System

Windows has one serious competitor: the Apple Macintosh operating system, usually referred to as (the) Mac OS. The Mac OS runs on the same computer hardware as Windows, but it is sold only with computer hardware manufactured by Apple Computer, Inc. (Microsoft does not build computers, and simply sells its operating system for use on any desktop computer hardware.) The Mac OS is used on about five percent of desktop computers, placing it in a distant but significant second place after Windows.

The Mac OS exists in two families, one of which is now obsolete and partially incompatible with the current family, Mac OS X. Earlier versions of the Mac OS required special hardware; the current version runs on the same desktop hardware as that used by Windows, although Apple does not sell the Mac OS for any hardware except its own.

An advantage of the Mac OS is its ease of use, which is legendary. Since the Mac OS is sold only with hardware built by Apple, the entire system functions as a unit with minimal problems. One

can simply turn on the computer and work. Many users prefer this simplicity to the slightly greater complexity of Windows, which can be installed on any hardware and thus may or may not be as simple to use as the Mac OS. Another advantage of the Mac OS is its traditional superiority for use with publishing, graphics, and music application software; this superiority has considerably waned and today Windows and the Mac OS are fairly equal in capability in these domains, but many artists still feel an emotional attachment to the Mac.

A disadvantage of the Mac OS is that it is available only for computers manufactured by Apple, and these computers tend to be much more expensive than other desktop computers. Another disadvantage is the much smaller number of applications available for the Mac OS. The Mac OS is a more closed system because of its tight association with Apple hardware, and less customization of the Mac environment is possible for people who like to "tweak" their computer systems—as a result, Windows tends to be popular with computer nerds, whereas the Mac OS tends to be popular with people who use computers only as tools and don't wish to be bothered with technical details.

Linux

Linux is a family of operating systems patterned after the venerable UNIX family of operating systems. It is used primarily on large server computers, but a small minority of desktop computer users run Linux as their operating system. Linux is used on considerably less than one percent of all desktop computers, and the users tend to be computer hobbyists.

Linux, like the UNIX operating systems that it emulates, is essentially a multiuser timesharing system, of a type popular in the 1970s. UNIX systems were designed to support hundreds or thousands of individual users simultaneously on a single large computer, via connections to individual *terminals* (screens and keyboards). Linux adapts this OS design to the desktop environment, with moderate success.

Linux exists in hundreds of different versions, called *distributions.* Each version contains the same core components, combined with a highly variable set of other system software that forms a more or less complete operating system. Some distributions are designed specifically for desktop use, others are not.

The desktop-oriented distributions of Linux provide graphic user interfaces reminiscent of those used on the Mac or Windows. The graphic interfaces vary considerably from one distribution of Linux to another, and in fact there are several different GUIs included within a distribution in some cases. The Linux family of operating systems allows considerable mixing and matching of components, such that each desktop computer running the system can have a version of Linux slightly different from any other. Hobbyists enjoy this, but it can be frustrating for ordinary computer users.

A significant advantage of the Linux family of operating systems is that they are usually available for free, at least in some form—although most Linux users buy their distributions from commercial distributors. They are also completely modifiable by the user, provided that the user knows a great deal about computers.

A serious disadvantage of Linux is that virtually no desktop applications are available for the operating system; and most of the popular commercial desktop applications that are the most widely on Windows and the Mac are unavailable for Linux. Another disadvantage of Linux is a lack of standardization—there are hundreds of different "flavors" of Linux available—and the need to have a very good knowledge of computers in order to install and use the operating system. These and other disadvantages of Linux have kept it from ever entering the mainstream of desktop operating systems, and will probably continue to do so for the foreseeable future.

Other Operating Systems

While there are many other operating systems for desktop computers, most are only used on a handful of machines, for very special purposes, or for experimentation. Many are obsolete, and some never were popular in the first place. The number and types of operating systems that can be designed for computers are limited only by the imagination, so it's reasonable to assume that there will be a continuing pattern of gradual change in computer operating systems in the future.