分析化学数据解析中的微机应用

FENXI HUAXUE SHUJU JIEXI ZHONG DE WEIJI YINGYONG

邵利民 编著

中国科学技术大学出版社

分析化学数据解析中的微机应用

FENXI HUAXUE SHUJU JIEXI ZHONG DE WEIJI YINGYONG

邵利民 编著

中国科学技术大学出版社

内 容 简 介

本书深入地介绍了一些常见化学计量学方法在分析化学复杂数据解析中的应用，包括傅里叶变换、曲线拟合、迭代法、插值、多元回归分析、主成分分析和化学因子分析等；通过数学原理的简要介绍，结合数据解析实例，阐明了各种化学计量学方法的性质、特点以及适用范围。此外，为方便研究者提高数据解析效率，本书还选择性地介绍了如何高效地使用微机操作系统以及 MATLAB 语言。

本书可作为分析化学专业研究生和高年级本科生的教材，也可供相关领域的教师和科研人员参考。

图书在版编目(CIP)数据

分析化学数据解析中的微机应用/邵利民编著. —合肥：中国科学技术大学出版社，2011.7

ISBN 978-7-312-02859-5

Ⅰ. 分… Ⅱ. 邵… Ⅲ. 计算机应用—分析化学—数据处理 Ⅳ. O65-39

中国版本图书馆 CIP 数据核字(2011)第 107027 号

出版 中国科学技术大学出版社

地址 安徽省合肥市金寨路 96 号，230026

网址 http://press.ustc.edu.cn

印刷 合肥现代印务有限公司

发行 中国科学技术大学出版社

经销 全国新华书店

开本 710 mm×960 mm 1/16

印张 10.5

字数 199 千

版次 2011 年 7 月第 1 版

印次 2011 年 7 月第 1 次印刷

印数 1—3000 册

定价 19.00 元

前　言

分析化学中的微机应用正在从常规性文档处理和资料获取等逐渐扩展到复杂实验数据的解析。这既是计算机软硬件性能提高的结果，也是当代分析化学发展的要求。

分析化学学科发展迅速，通过对仪器分析手段不断地改进和创新，已经能够对复杂化学体系的多种化学性质进行测量，而且可以在短时间内测得大量数据。这些实验数据具有量大、信息丰富的优点，但同时也带来诸如结构复杂、干扰严重以及冗余信息多的问题。可以说，在当代分析化学中，测量到的信息不一定是可见的。在应用方面，如食品、医药、环境以及国土安全等，不仅要求分析仪器可靠、便携，而且要求数据分析快速、准确。在这种背景下，如果仍采用传统的简单数据分析方法，就无法充分发掘和利用包含在原始数据中的信息。因此，当代分析化学研究需要高级数据解析方法，这也是化学计量学——以数学和统计学方法研究化学问题的学科——日益得到重视的原因之一。

化学计量学的广泛应用逐渐凸显了所谓的“黑箱”问题。一方面，数量众多、界面友好的工具软件极大地方便了高级数据解析方法的应用，但是研究人员使用这些软件时，绝大多数情况下只能关注实验数据的输入和计算结果的输出。另一方面，数据解析方法中较复杂的数学部分让人却步。这种只知道输入输出而不清楚方法性质的“黑箱”问题导致了高级数据解析方法的误用和滥用。对化学计量学而言，研究工作不仅仅是如何使用方法，更重要的是如何对数据解析结果进行阐释；越是高级、复杂的解析方法，我们越应当慎重对待其解析结果，因为计算出的结果不一定是真实的。为了避免“黑箱”问题，有必要深入理解这些高级数据解析方法的原理和性质，唯此，我们才能够选择合适的方法，较全面地发掘实验数据中的显性和隐性信息，合理地阐释数据分析结果，进一步提高数据解析的质量和结论的可靠性。此外，化学计量学专业人员只有深入掌握了方法的基本数学原理，才有可能进行创新。基于这样的事实，本书在介绍高级数据解析方法时涉及了较多的数学

知识，并且强调了解析方法的纯数学结果与化学组分的具体化学信息之间的关系，这也是本书的指导思想。

本书的主要内容是一些发展相对成熟的高级数据解析方法，如基于傅里叶变换的滤噪和信息提取、非线性最小二乘曲线拟合、多元回归分析、用于分析信号处理的主成分分析等等。考虑到分析化学专业人员的数学知识背景，本书对一些方法的数学原理和推导过程进行了必要的精简和重组，在确保严密性和完备性的前提下，做到尽量简化。另外，编著者结合自己的科研工作，选择了一些有代表性的数据分析实例说明这些数据解析方法的应用，期望通过这种方式，使读者能够在更高层次上理解这些方法的性质、特点和适用范围，同时又尽量避免其中相对难懂的数学部分可能对学习积极性造成的挫伤。为了提升分析化学研究中微机应用的可操作性和舒适度，从而最终提高工作效率和生产力，本书还简要介绍了计算机硬件方面的知识，包括模/数、数/模转换；并介绍了计算机应用方面的知识，包括操作系统和主流数据处理软件 MATLAB 语言的高效使用。

显然，除本书涉及的内容之外还有很多优秀的数据解析方法，但本质上方法只是“用”，正确的应用以及对结果合理的阐释才是“体”，这也是编著者期望通过本书传达的一个观点。

化学计量学是一门较新的交叉学科。一方面，其优点正逐渐为广大分析化学工作者所了解和接受；另一方面，在研究中应用化学计量学方法有一些“时尚”味道。编著者奢望通过本书降低数学方法在分析化学数据解析中的神秘性，减少只为“时尚”而研究的现象，就像兰(A. Lang)讽刺一些人对统计学方法的使用：He uses statistics as a drunken man uses lamp-post ... for support rather than illumination.

本书从构思到完成，历经数年。很多人提供了资料、帮助和支持，这里一并表示感谢。同时，本人感觉努力写好本书是对他们最好的报答。编著者深知自己水平有限，所以始终小心谨慎，在材料取舍、公式推导等方面都尽自己最大的努力。尽管如此，疏漏在所难免，恳请读者不吝赐教，以使本书内容更趋完善。最后，通过本书，我想表达一个愿望：希望我的儿子尽他所能，以认真、深刻的态度对待自己的工作和生活。

邵利民

2011 年 1 月于中国科学技术大学

目　　次

第一章　高效使用 MS Windows 操作系统

操作系统是一种控制计算机的软件，可以协调计算机内部的各项活动，提供运行程序和管理数据等多种服务。操作系统是应用软件或计算机用户与计算机硬件之间的一个中间层，也是一个管理层，使对硬件的操作更加方便和安全。操作系统还提供了另一层次的功能，就是协助计算机完成它的基本硬件操作。操作系统，特别是界面友好的操作系统，在微机的普及方面有重要作用。一般认为，如果需要某个软件的唯一原因是用户拥有一台计算机，那么该软件就是系统软件。

第一节　MS Windows 的基本概念

Windows 操作系统是基于图形用户界面（GUI）的系统，加上微软公司在用户界面（UI）和用户体验（UX）方面的努力，即使是初级用户也能基本正常使用 Windows 操作系统。然而，Windows 操作系统的功能却远不是那么一目了然。对普通用户来说，对操作系统进一步的了解有助于提高用户体验和工作效率。

一、分区（Partition）和卷（Volume）

分区是对物理硬盘的划分。在 MS 操作系统中，主分区（Primary）是操作系统所在的分区，主分区分配盘符 C（由于历史原因，字母 A 和 B 被分配给软盘驱动器）。扩展分区（Extended）从属于主分区，而且只能有一个。在扩展分区上可以建立多于一个的逻辑驱动器（Logical Drive），并且被分配给不同的盘符，是 C 之后的字母。“卷”和逻辑驱动器在概念上是等价的，而且和文件系统密切相关。

显然在同一硬盘上建立多个逻辑驱动器，或者使用多个物理硬盘，能够高效、

灵活地组织各种文件。需要说明的是，迄今为止没有来自微软的官方证据表明将操作系统和应用软件安装在不同的逻辑分区，甚至是不同的物理硬盘上会对系统性能产生明显影响。

二、文件系统(File System)

文件系统是操作系统管理文件的方式，在 MS 操作系统中，文件系统的架构是文件分配表（File Allocation Table，FAT）和 NTFS（New Technology File System）。基于 FAT 的文件系统的发展经历了不同的版本，如 FAT12、FAT16、FAT32。NTFS 是较 FAT 更先进的文件系统，在目前主流操作系统中被广泛应用，如Windows XP 和 Windows 7 等。

关于文件系统的一个有趣的例子是 USB 移动存储设备。出于兼容性考虑，绝大多数 USB 移动存储设备都被预格式化为 FAT 文件系统，而不是 FAT32 或 NTFS。FAT 系统的根目录最多可容纳 512 个条目，而非“8.3”格式文件名的文件至少要占用两个条目。结果是，FAT 格式的 USB 移动存储器上，根目录下最多能容纳 512 个文件。如果复制文件超过此限额，将出现错误“The directory or file cannot be created”。解决办法很简单：在根目录下建立子目录，将文件分放在不同的子目录中；或者将 USB 移动存储设备格式化为 FAT32 或 NTFS 文件系统。对于如何将 USB 移动存储设备格式化为 NTFS 系统，Windows XP 不提供直接选项，但用户可以在 USB 移动存储设备的右键菜单中选择“属性”，在“硬件”选项卡中选中该 USB 移动存储设备，并点击“属性”按钮；在“策略”选项卡上，选中“为提高性能而优化”。然后，在该 USB 移动存储设备的“格式化”对话框中，就会出现 NTFS 选项。

三、窗口(Window)

在图形用户界面的操作系统或基于该系统的应用程序中，窗口是屏幕显示区域的一部分（有时可以是全部），窗口通常带有边框，有一致的外形（如边框、标题栏和按钮等）和通用的操作（如移动、最大化等），窗口包含自己的文档或消息。Windows 操作系统下的绝大部分应用程序都是通过窗口完成人机交互的。

对应于多任务，Windows 允许同时存在多个窗口。但任一时刻只能有一个窗口可以响应用户操作，进行人机交互。这样的窗口称为活动窗口（Active Window），其余窗口称为非活动窗口（Inactive Window）。活动窗口和非活动窗口的切换可以通过任务栏上的按钮或“Alt + Tab”来进行。从放置的顺序来看，一个 Window 可以是 top、topmost、non-topmost。Windows 的这一属性被一个称为 Z-order 的堆栈管理。具有 top 属性的窗口被置于 Z-order 的顶端，遮盖了其他

non-topmost 窗口，但低于具有 topmost 属性的窗口。一般地，一个 active 窗口不一定必须是 topmost；反之亦然。

一般来说，将某进程的窗口最小化可以减小其对资源的占用，因为系统不需要进行窗口重绘。

四、文件(File)

文件是一个存储在一定介质（比如磁盘）中完整的、有名称的信息的集合。文件包含的信息可以是可执行代码、数据、音频、视频，或者其他任何信息。文件的存在使得计算机能够把一组信息与另一组信息分开。文件存储的基本单位是簇（Cluster），每个簇由一组扇区（Sector）组成。在文件系统上，一个文件的真实大小不一定是一个簇的整数倍，因此文件大小一般不等于其占用的磁盘空间。显然，簇小的文件系统能更有效地利用磁盘空间。

文件通常采用“文件名.扩展名”的方式命名。文件名说明文件的内容，扩展名则指定了这个文件的类型，比如“TXT”是文本文件，“EXE”是可执行文件，“DOC”是 MS Word 文档。需要注意的是，扩展名不决定文件的类型，只起说明作用。在 MS DOS 下，文件名和扩展名分别只能有 8 个和 3 个字符，即所谓的“8.3”格式，这样文件的命名就有很多限制，也有很多不便。在 MS Windows 操作系统中，允许文件名长度可达 255 个字符，而且也可以有“空格”、“.”等字符，这样，文件名就可较完整地表达文档内容，方便文件管理。文件命名一般采用简洁、清楚的原则。

在 Windows 操作系统中，对文件的操作，如“复制”、“移动”、“删除”和“重命名”等可以通过鼠标右键菜单中的对应项或快捷键来完成，在 DOS 中则使用命令“copy”、“move”、“del”和“ren”。DOS 下还有一个命令是“列出文件目录”：“dir”。

五、文件夹(Folder/Directory)

文件夹也称子目录，是操作系统对磁盘文件的组织管理方式。文件夹是一个形象的称呼，通常情况下包含同一类的文件，并以一个确定的名字（和文件不同，文件夹多数不使用扩展名）来说明，比如“Backup of currently used software”。从技术的角度来看，文件夹是一种特殊类型的文件，是包含了一个文件列表的文件，这个文件列表中的文件就称在这个文件夹中。

在 Windows 操作系统中，打开一个文件夹（通过双击该文件夹的图标）就会出现一个窗口，这个窗口显示了包含在该文件夹中文件的名称、大小、类型以及修改时间。在 DOS 中，对文件夹（此时称子目录更合适一些）的类似操作是通过命令“cd”来完成的。

在 Windows 操作系统中，对文件夹的操作和对文件的操作非常类似，也可以通过右键菜单中的对应项实现“复制”、“移动”、“删除”或“重命名”。DOS 中对子目录的操作稍微复杂一些，而且没有“重命名”命令，但我们可以借助外部命令“move”来完成。

需要注意的是，Windows 操作系统中还有一些特殊的文件夹，如“控制面板”、“拨号网络”以及“打印机”等，属于系统文件夹，一般不允许用户修改。

六、文件/文件夹属性(Attribute)

为了便于管理，操作系统中对文件和文件夹规定了四种属性，分别是“只读”、“存档”、“隐藏”和“系统”(分别对应 DOS 下的 r、a、h、s 属性)。在 Windows XP 中可以通过鼠标右键菜单中的“属性”项对文件或文件夹的属性进行修改，但通常只能是“只读”和“隐藏”，不能修改“系统”属性，除非借助其他工具或在 DOS 下使用“attrib”命令。所以，在微机操作易用性为主的时代，DOS 命令依然有用武之地。下面是利用 DOS 命令清除一个恶意程序的例子。

在注册表的自启动位置 HKEY_LOCAL_MACHINE\Software\Microsoft\Windows\CurrentVersion\Run 下发现一个奇怪的程序：XP-14FF44BE.EXE，见图 1-1(a)。而在启动文件夹 Startup 下发现一个不显示名称的快捷方式，见图 1-1(b)，查看其属性，指向的就是文件 XP-14FF44BE.EXE。在文件夹 system32 下，用命令“attrib”查看 XP-14FF44BE.EXE 的属性，发现是“系统”、“隐藏”和“只读”，见图 1-1(c)。恶意软件无疑！使用命令“attrib”使之现形，删除即可，见图 1-1(c)。

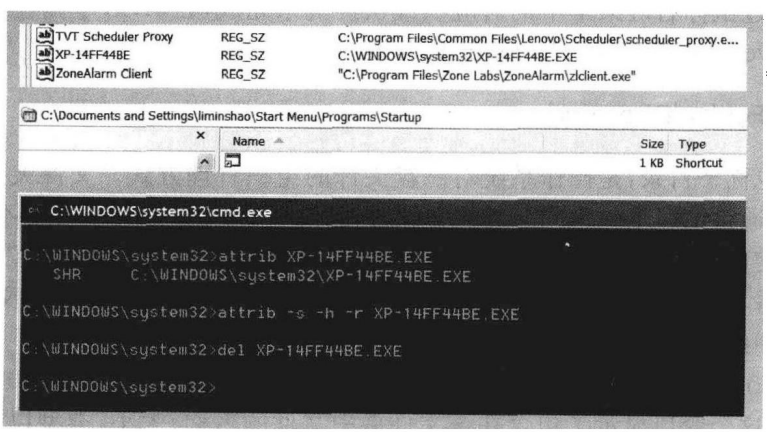

图 1-1　使用 DOS 命令清除 Windows XP 中一个恶意软件

缺省设置下，Windows 操作系统不显示具有“隐藏”或“系统”属性的文件和文件夹，以防止用户的误操作。如果要显示这些文件，可以通过“开始”→“设置”→“文件夹选项”（或者资源管理器“查看”菜单中的“文件夹选项”），在“文件夹选项”窗口中的“查看”选项卡进行设置。在 DOS 方式下，使用命令“dir/a”。

七、时间戳(Timestamp)

时间戳是操作系统对文件或文件夹时间属性的设置，是文件（夹）的重要信息之一，特别是在跨时区协作或取证等场合。Windows 对时间戳的处理和文件系统有关：在 FAT 中，时间戳根据本地时间设置，而在 NTFS 中，时间戳根据 UTC（世界标准时）设置，显示时转换为当地时间，因此不受时区或夏令时的影响。

Windows 操作系统中的时间戳有三个，分别是“创建时间”（Created）、“修改时间”（Modified）和“访问时间”（Accessed）。“创建时间”是一个文件（夹）被创建的时间，“创建”指的是产生新文件（夹）或者复制已有文件（夹），但不包括对已有文件（夹）的移动或重命名。文件的“修改时间”是其内容最近一次发生改变的时间；文件夹的“修改时间”则是指位于其中的文件最后一次被创建、复制或重命名的时间，但不受文件编辑的影响。显然，文件（夹）的创建时间可以后于修改时间。“访问时间”是文件（夹）最近一次被访问的时间。

八、快捷方式(Shortcut)

快捷方式是一种快捷方式类型的文件，其扩展名为“lnk”。快捷方式相当于一个索引，指向一个文件或文件夹，访问某个快捷方式就相当于访问该快捷方式指向的文件或文件夹。

在不同位置放置同一对象的多个副本既浪费空间又容易出现失误——当其中一个副本更新时，很可能会忘记更新其他副本。而快捷方式则可以解决这一问题：首先一个快捷方式的文件很小；其次放在不同位置的快捷方式都指向同一个对象；另外可以创建同一对象的多个快捷方式。打开快捷方式意味着打开其所指向的对象，删除快捷方式却不影响对应的对象实体。

使用鼠标右键菜单非常容易地建立快捷方式：可以通过“新建”→“快捷方式”；也可以先复制对象文件，然后选择右键菜单中的“粘贴快捷方式”。一个快捷方式及其所指向的对象使用同样的图标，只是左下角多了一个小箭头，以示区别。

快捷方式图标中的小箭头可以换成其他图形，甚至可以去掉。有一些应用程序也提供了这样的功能。显然，这种做法没有任何积极的意义。

九、硬链接(Hardlink)

硬链接是 NTFS 文件系统(不能用于 FAT 文件系统)中对一个文件的链接,可以理解为对实体文件的一个索引。在 NTFS 中,一个文件可以有多个硬链接,但不能超过 1 023 个;硬链接必须位于同一卷中,即不能在 C 盘上创建 D 盘文件的硬链接。只有当一个文件的所有硬链接都被删除后,该文件占用的空间才被标记为可用,也就是该文件被删除了。

硬链接和快捷方式有些相似,但本质是不同的。例如有一个文本文件 Example.txt,我们创建该文件的一个快捷方式和一个硬链接。快捷方式是一个不等同于源文件、但指向 Example.txt 的 lnk 文件,而硬链接就是 Example.txt 本身。在 Windows "资源管理器"中,快捷方式是个小文件,一般不等于其指向的文件的大小;硬链接的大小就是文件的大小,但并不额外占用磁盘空间。比如,我们为 1 MiB* 的一个文件创建了 1 000 个硬链接,在"资源管理器"中,这 1 000 个硬链接都显示1 MiB 的大小,但总的实际占用空间只有 1 MiB,而不是 1 000×1 MiB≈1 GiB。

硬链接可以视为实际文件,而又不额外占用磁盘空间,所以在某些场合下特别有用。比如,需要备份的文件分散在不同的文件夹中时,逐个备份费时费力而且容易出错,在 NTFS 系统中可以创建这些文件的硬链接,然后放在一个特定的文件夹中,只备份该文件夹即可。

在 Windows XP 中,使用命令"fsutil"创建硬链接。

借助硬链接的备份无法应用于 MS Word 文档,因为 Word 文档修改后保存时,MS Word 实际上先把文档内容写入一个临时文件,然后删除旧文件,再把临时文件重命名为源文件,同时设置时间等信息。结果就断开了原来文档的硬链接。

十、关联(Associate)

操作系统根据某个特定的文件扩展名将具有该扩展名的文件与某个应用程序相联系,当打开这类文件(通常是双击鼠标左键,或者选定文件后,敲击"Enter"键)时,操作系统会自动启动那个应用程序来打开文件。

关联的重要作用是简化操作,尽管我们完全可以先启动应用程序,然后在应用

* 1 MiB=1 024 KiB,1 KiB=1 024 Byte。注意这与 MB、KB 不同。

程序中打开要处理的文件。关联不需要一一对应：某个应用程序可以关联到多种类型的文件；而某种类型的文件也可以有多个打开方式（当然只能有一个默认）。

被关联的文件对象通常被指定一个特殊的图标，而没有被关联的则被系统分配一个统一的图标。所以在资源管理器中很容易识别出某类文件是否关联。

关联的建立可以通过鼠标左键双击某一未知类型的文件，然后在“打开方式”窗口中选择某一应用程序。关联的编辑可以在“文件夹选项”窗口中的“文件类型”选项卡进行（资源管理器“查看”菜单中的“文件夹选项”）。也可以选定某文件后，在鼠标右键菜单中选择“打开方式”。在 MS DOS 方式下，应用关联的命令是“start”。

十一、拖放（Drag-and-drop）

拖放是图形用户界面操作系统特有的操作方式，通过鼠标拖动对象的方式完成操作。比如将一个文档用鼠标拖动到“回收站”图标上释放，就会删除该文档；如果拖动到“打印机”图标上释放，就会打印该文档。再如，将一个文本文件（也可以是其他任何类型的文件）拖到“记事本”（Notepad）的窗口，“记事本”就会显示该文件的内容，就相当于用“记事本”打开该文件（但对于 MS Word，拖放一个非 Word 文件是在当前文档中建立一个对象包，而不是打开该文件）。“拖放”也可以用于文件的复制（同时按“Ctrl”键）、移动（同时按“Shift”键）和建立快捷方式（同时按“Alt”键）。“拖放”的优点是操作形象简单，但也容易导致误操作。一个小技巧如下：如果拖放失误造成了不必要的文件复制或移动，可以通过“Ctrl + Z”快捷键还原这些误操作。

十二、回收站（Recycle Bin）

回收站是 Windows 9x 系统开始引入的一个重要功能。用户删除的文件首先被放入“回收站”，需要时能够从“回收站”恢复。“回收站”使用一定的磁盘空间来存放被删除的文件，用户可以指定这个空间的大小，所以如果被删除文件太大，就不能放入“回收站”，此时系统建议“直接删除”。如果用户要真正删除，就要选择右键菜单中的“清空回收站”或使用“Shift + Delete”删除文件。除非借助特定的工具，被“清空回收站”命令或“Shift + Delete”删除的文件一般不能恢复。

十三、任务（Task）

一个独立的应用程序或可以作为单独实体独立运行的子程序处于运行状态就称为任务。在 MS Windows 操作系统可以同时运行多个任务。查看任务最直接的方式是查看任务栏上的图标，但有些程序不在任务栏上显示图标，比如一个时钟程序，这时可以在“Ctrl + Alt + Del”激活的“关闭程序”窗口中看到正在运行的任务。

还有一些任务使用上面的方法仍然看不到，一般来说，这是操作系统正在使用的各种服务程序，不对用户透明是为了避免不必要的失误造成系统崩溃，如果不属于这一类，那几乎可以肯定是病毒或木马。

在多任务状态下，任何时刻只能有一个任务具有系统的控制权并对用户发出的指令做出响应，这称为前台(Foreground)，前台任务的窗口就是“活动窗口”；与之相对应的就是不需要用户的交互而正在进行的任务，称为后台(Background)，后台任务的窗口是“非活动窗口”。后台任务与前台任务相比所分配的优先级较低，并且它们通常对用户是不可见的，除非用户请求一个更新或将该任务拉至前台。不同任务的切换可以通过点击任务栏上对应于不同任务的图标，或者通过“Alt + Tab”组合键。

从用户角度来说，任务就是程序；从 CPU 和系统资源分配和调度的角度来说，任务也称为进程(Process)。进程是一个具有一定独立功能的程序在一个数据集合上的一次动态执行过程。进程与处理器、存储器和外设等资源的分配和回收相对应，是计算机系统资源的使用主体。所谓多任务处理，实质上就是操作系统中多个进程的并发执行。操作系统是以进程为单位对资源进行分配和调度的。进程和程序的区别如下：

(1) 进程是程序的执行过程，是动态的过程，属于一种动态概念。程序是一组有序静态指令和数据的集合，用来指示处理机的操作，是一种静态概念。

(2) 从结构上看，每个进程实体是由程序段和相应的数据段两部分构成的，并且进程结构中还要包含 PCB，即进程控制块。

(3) 一个进程可以涉及一个或几个程序的执行；反之，同一程序可以对应多个进程，即同一个程序段可以在不同数据集合上运行，构成不同的进程。

(4) 进程能真实地描述并发执行的过程，而程序仅仅是静态指令堆积的序列。

(5) 进程有创建其他进程的功能，而一般的程序不具有创建其他程序的功能。

(6) 每一个程序都是在一个进程现场中运行的。

在只有进程概念的操作系统中，进程是存储器、外设等资源的分配单位，同时也是处理器调度的对象。为了提高进程内的并发性，进一步提高处理器的利用效率，Windows 98 引入了线程(Thread)的概念，把线程作为处理器调度的对象，而把进程作为资源分配的单位。线程也称为轻量级进程(Lightweight Process)，它表示进程中的一个控制点和一条执行途径，任何进程都可以创建多个并发执行的线程。由于同一进程内各线程都可访问整个进程的所有资源，因此它们之间的通信比进程间的相互通信要方便，而同一进程内的线程间切换也会由于许多上下文的相同之处而得到简化。

十四、虚拟内存(Virtual Memory)

虚拟内存是 Windows 管理所有可用内存的方式。对于 32 位系统，每个进程可用的内存地址从 0 到 $2^{32}-1$。所以每个进程最大可占用 4 GiB 的内存；其中 2 GiB 是与操作系统以及其他所有进程共享，另外 2 GiB 为该进程所独占(这就是常说的 32 位 Windows 中一个进程最多能用 2 G 内存的由来。)

在这 4 GiB 虚拟地址空间(Virtual Address Space，VAS)中，独占的 2 GiB 是进程隔离的，比如三个进程最多能用 2 GiB(共享) + 3 × 2 GiB(独占) = 8 GiB。Windows XP 最大只支持 4 GiB 的 RAM，因此需要页面文件(pagefile.sys)来补充不足。此 4 GiB 的虚拟地址空间按照 4 KiB 的大小进行分页(Page)，然后以页为单位映射到实际存储单元中，包括物理内存(RAM)、页面文件(Page File，在 Windows 9x 中称为交换文件(Swap File)，即 win386.swp)和其他文件自身(比如一些长时间未活动的进程的 exe 文件本身)。可见，通常人们所说的虚拟内存实际上只是指其中的页面/交换文件而已。

RAM 除了保存最近读写的文件缓存外(File Cache，相当于 Windows 9x 中的 Vcache)，主要用来存储正在使用的程序代码和数据，当 RAM 资源紧张，或者有程序代码或数据长时间未使用时，XP 通常会将非活跃程序代码所在的地址页映射回程序文件(exe、dll 等文件)，将数据所在的地址页映射到页面文件中并复制数据，然后将它们本来占用的 RAM 空间释放。这个过程称为页出(Page Out)。当某进程访问某个内存地址，而该地址所在的页不在 RAM 中时，将产生一个页面错误(Page Fault)中断，告诉系统从页面文件或程序文件中取回包含该地址的虚拟内存页，即将内容复制到 RAM 并建立新的虚拟地址映射，并将页面文件中对应部分标记为未使用，这个过程就是页入(Page In)。页入成功的话，结果就是一个 Valid Page Fault，否则就是 Invalid Page Fault。前者非常普遍(可以在任务管理器的进程页监视到)，而后者是由程序或硬件错误引起的，如果发生在进程上会导致非法操作，如果是系统本身则很可能发生蓝屏现象。

十五、快捷键(Keyboard Shortcuts)

快捷键是指图形用户界面中，能够完成一个鼠标操作功能的按键，通常是组合键。即使现在鼠标的作用非常大，但我们仍不能脱离键盘，快捷键有时能比鼠标操作更加迅速、准确，而且在鼠标失灵时(尽管很不常见)能够正确完成操作。比如使用“Ctrl + X”、“Ctrl + C”和“Ctrl + V”分别来完成“剪切”、“复制”和“粘贴”，通常比用鼠标点击相应的按钮要快得多。

MS Windows 操作系统为各种操作定义了众多的快捷键，只要记住一些就够了，比如“Win + E”（打开资源管理器）、“Win + D”（显示桌面）、“Alt + F4”（关闭窗口/程序）、“Esc”（取消）、“Enter”（确定）、“Delete”（放到“回收站”）、“Shift + Delete”（直接删除）和“F2”（重命名）。

十六、人机界面（Shell）

操作系统所进行的工作远不只是我们在显示器上看到的那样。操作系统通过所谓的核心组件（Kernel）忙于和各种硬件打交道，包括调配资源和管理输入输出，同时还要“煞费苦心”地将执行结果“翻译”成我们能理解的形式。“翻译”就是通过人机界面进行交互，所需要的程序统称为 Shell。在 MS Windows 操作系统中，Shell 提供了各种可视化组件，包括窗口、按钮、对话框等。

操作系统并不需要 Shell。事实上，如果没有 Shell，操作系统的负担还轻些。作为用户，我们须臾离不开 Shell，而且希望（要求）界面更美观、更友好。下面以 Windows XP 为例，进一步了解 Shell。

“桌面”对应于文件夹“C:\Documents and Settings\%user%\Desktop”，或者“C:\Documents and Settings\All Users\Desktop”。“桌面”上的各图标就是文件夹中文件的图标。在此文件夹下新增一个文件，其图标就会出现在“桌面”上；如果我们在“桌面”上创建一个快捷方式，文件夹中就会出现一个快捷方式类型的文件。（对多用户功能较弱的 Windows 98 来说，路径则是 C:\Windows\Desktop。）所以我们既可以通过桌面上鼠标右键菜单，也可以通过操作上述文件夹中的文件达到对桌面快捷方式的定制。

“我的文档”对应于 C:\Documents and Settings\%user%\My Documents，下面通常会有另外一些文件夹，如“My Music”和“My Pictures”等，即中文版操作系统中的“我的音乐”和“我的图片”等。

“回收站”对应于“X:\RECYCLER”（X 表示相应的盘符）。这是一个具有“隐藏”和“系统”属性的文件夹。被放入“回收站”的文件就位于“X:\RECYCLER”下的相应子文件夹中。用户可以通过图形界面方式，也可以在 MS DOS 方式下将文件复原。

Windows XP 的“快速启动工具栏”对应于文件夹“C:\Documents and Settings\%user%\Application Data\Microsoft\Internet Explorer\Quick Launch”。

图形用户界面的 Shell 极大地增强了操作系统的易用性。但缺点是执行效率和功能方面的削弱。当然，对普通用户来说，这不是问题；对高级用户来说，这不成问题。

我们在 Windows XP 上做个小实验（恶作剧），感受一下 Shell 的重要性。打