

97,24(6)

Java

程守仁

计算机网络

①

当前 Java 发展综述

An Overview on the Current Development of Java Language

汪审权 胡守仁

(国防科技大学计算机学院 长沙 410073)

1-104

TP312Ja

TP393

1-5

摘要 The paper describes the main development direction of Java Language and Java Technology. Analyses the importance of studying Java Language and Java Technology in our country.

关键词 Java, Applet, Java Virtual Machine, Network Computers, WWW, Internet, Intranet, Extranet

Java 由一个面临失败工程的产物成为全世界炙手可热的语言,只有短短的两三年时间。现在,除了 Java 语言最初的开发者 SUN 公司成立专门的公司 JavaSoft 来研究和发 展 Java 语言和基于 Java 的各种产品以外,世界著名的其它大公司如:IBM, Netscape, Intel, SGI, Oracle, Microsoft 等都对 Java 语言表示了极大的兴趣,投入了大量的人力、物力来进行 Java 语言及其相关技术的研究和开发。1995 年和 1996 年是 Java 年,除了计算机专业人员以外,还有原来与计算机联系不太紧密的商家和公众舆论也对 Java 表示了浓厚的兴趣。这是前所未有的。这两年,Java 是如此之热,以至于在报刊杂志和 Internet 上随处都可以看到 Java。许多人,包括许多大公司的负责人,在谈论计算机的发展时,言必称 Java。

确实,谁都不可能想到,只有短短的两年多一点的时间,Java 就会变得如此来势汹涌,仿佛整个计算机界都已经进入到 Java 的汪洋大海之中。在当前 Java 发展如此迅速的情况下,我们迫切需要对 Java 及其相关技术有一个较为清醒的认识,以顺应计算机技术发展潮流。为此,我们不妨从 Java 产生的背景及它火热起来的原因入手,分析当前 Java 及其相关技术的发展现状,并在此基础上讨论 Java 技术给我们带来的机遇和挑战。

一、Java 发展的背景和原因

八十年代以前,计算机的研究和应用大都局限于一个个独立的计算机系统,各计算机公司所销售的计算机系统的体系结构、系统软件各不相同,因此,用户须针对特定的计算机系统编制应用软件,软件的兼容性问题一直困扰着人们。随着计算机网络

的发展,用户希望各种类型的计算机系统能够通过网络互连,达到数据通信、数据共享、资源共享。尽管通过采用统一的通信协议和转换接口,可以部分地实现这些目的,但是由于不同的计算机系统之间的巨大差异,使得人们很难在不同的计算机系统之间进行真正自由的相互访问。为了解决这个问题,人们提出了开放式系统的概念,企图通过统一的计算机接口标准,解决平台的相关性,使计算机网络中的异构计算机系统之间能够实现资源共享和互操作。开放式系统强调可互操作性、规模可变性 (scalability)、可移植性和可连接性,对于绝大多数的公司和用户来说,确实是一个十分美好的想法。但是,由于技术上和商业上的原因,开放式系统并没有发展到人们所期望的程度。几乎所有应用程序的开发还是基于某个特定的计算机平台,不同的系统平台之间不得不做大量的移植工作。随着应用软件的规模越来越大,这种移植工作也变得越来越困难。在这种情况下,人们对实现计算机平台无关性的愿望越来越强烈。

计算机网络的出现,使计算机的应用和发展呈现出一片极为广阔的天地。但同时,也使人们深切感受到现有的软、硬件系统不能适应当前网络环境下的程序设计和开发的需要。人们迫切需要有一种新的语言、新的技术、关于软件开发的新思想,使人们在网络环境下开发程序的时候,不用去考虑这个程序将在何种系统下使用,以充分发挥计算机网络的优 势。Java 在这个时候的出现,它所宣称的平台无关性和对网络编程的强大支持能力,对计算机界来说,真可谓是“久旱逢甘露”。当然 Java 能变得如此红火,还有以下的几个因素:

1-10 1.

首先,是 Internet 发展的需要。Internet 已经发展了好些年,但在 Java 和 WWW 出现以前,Internet 的发展并不很快,它的使用范围也仅局限于一些科研院所的专业人员。在 Internet 上只能传递简单的文字信息。1991 年夏天,WWW 的出现马上使 Internet 的面貌发生了翻天覆地的变化:Internet 的使用大众化了。而 Java 语言的出现,则更为 WWW 和 Internet 的飞速发展装上了强有力的引擎。Java 不仅使 Web 变成动态的了,而且更重要的是 Java 使 Internet 在传输数据的同时,也能传输服务。

其次是软件产业界的激烈竞争和力图打破 Wintel 一统天下的局面的需要。传统的 PC 机市场几乎被 Intel 芯片和 Microsoft 软件所垄断,并且,随着计算机使用越来越广泛,Wintel 正在不断地扩大自己的垄断优势。其它的软件厂商虽然希望能改变这种对自己越来越不利的局面,但是在传统的 PC 机操作系统和应用软件的市场上,几乎已经不可能撼动 Wintel 的霸主地位。Java 的出现,使软件开发商们看到了打败 Microsoft 的一线曙光。从以往计算机发展的历史看,每一次计算机技术的重大突破,都会引起计算机产业界格局的变化和结构的重组。想当年 DEC 公司在小型机领域和 IBM 公司在大型机领域的王者地位,何其风光。IBM PC 的出现则给 Intel 和 Microsoft 的腾飞提供了契机。现在,Java 的出现,又使大家看到了可以进行放手一搏的绝好机会;SUN,IBM,Oracle,Netscape,SGI... 谁都希望能早日结束 Wintel 王朝,并在新一轮的市场竞争中处于领导者的地位。他们组成了 Java 联盟,竭力扩大 Java 的影响。当然,Intel 和 Microsoft 是不可能轻易放弃自己的王者地位的。他们也看到了 Java 的发展潜力,也希望把 Java 纳入到自己的发展轨道上来。不约而同地,大家都发起了 Java 攻势,积极展开对 Java 技术的研究和 Java 应用程序的开发,把 Java 炒得沸沸扬扬。

二、Java 语言及其系统的核心技术和应用的发展现状

为了把握 Java 的发展方向,我们考察了最近几个月以来 Java 及其相关技术的发展情况。发现目前 Java 的发展主要集中在以下的几个方面:

1. Java 核心技术的发展;
2. Java 开发工具的不断更新换代;
3. 基于网络的 Java 应用程序的迅速推出。

Java 的核心技术主要包括以下几个方面的内

容:Java 虚拟机(JVM:Java Virtual Machine)具有“即插即用”功能的 java beans 和号称“编写一次,到处可用(write once,run everywhere.)”的 javaOne。

Java 的开发工具影响较大的有:Java Soft 的 JDK (Java Development Kit); Symantec 公司的 Symantec Cafe Lite;Microsoft 的 VJ++;Borland 公司则把 Java 和 BC++5.0 捆绑在一起等等。

下面我们择要讨论一下上面的内容。

首先 Java 技术的核心是 Java 虚拟机 JVM。JVM 是 Java 程序的执行平台。Java 程序被编译成 JVM 上的“机器码”——字节码,由于字节码是与机器无关的,因此 JVM 较好地解决了跨平台的问题,但目前,Java 程序的执行速度还不令人满意。究其原因,主要是由于 JVM 中的“垃圾收集器”的存储管理不够强大,JVM 需要花费大量的时间对对象进行管理。用 Benchmark 进行测试可以发现,JVM 在普通指令如算术运算、逻辑运算、循环等的速度几乎与 C++不相上下,但是有关对象操作的指令则慢了一个数量级以上。由于 Java 是一种完全面向对象的语言,对对象操作的低效导致了整个程序运行速度的明显降低。Java Soft 公司也看到了这个比较致命的问题,并着手对 JVM 进行改进;他们收购了 LongView 技术公司,利用该公司的编译技术实现了一种新的虚拟机和垃圾回收技术,使 JVM 的性能提高了 10 倍。对 JVM 来说,其核心问题是提高 Java 程序的执行效率。Java Soft 公司最新推出的 JRE 1.1.1(The Java Runtime Environment,Version 1.1.1)是第一个公开的有关 Java 最小运行时间的版本。它仅包括 Java 虚拟机、Java 的核心类和一些支撑文件,但不包括编译器、调试器和工具,其目的也是为了提高 Java 程序的运行速度。当然,Java Soft 公司还在硬件上为提高 Java 程序的执行速度进行了两年努力;他们推出的 Java Station 就是一种 NC (Network Computer)。

其次,Java Soft 公司另一个引以自豪的技术是 Java Beans。Java Soft 给 Java Beans 的定义是:“A Java Bean is a reusable software component that can be manipulated visually in a builder tool”。也就是说,Java Beans 是一种软构件。虽然,Java 的与平台无关性本身就意味着 Java 程序可以“编写一次,到处可用”。但是,“编写一次,到处可用”并没有意味着程序中功能部件的可重用性。Java Beans 则是在“编写一次,到处可用”的基础上更进了一步,它强调部件的可重用性,并通过事件模型进行链接,希望能使

软部件也能象硬插件一样能实现“即插即用”，不仅使软件开发的速大为提高，成本大为降低，而且软件的质量也大为提高。因此，相信在可预见的将来，软构件将成为软件设计和开发的主要形式。

说到 Java Beans，就需要提一下与 Java Beans 具有相似概念的由 Microsoft 提出来的 ActiveX。两者都是软构件，在功能上具有相似之处。作为各自公司的王牌，两者都希望能超过对方。目前，Java Beans 所具有的优势是平台无关性，但 ActiveX 则具有更强的功能，能更好地利用 Windows 的各种资源。现在，开发者已经对小应用程序的简单功能感到不满。这迫使 JavaSoft 公司开始考虑在安全性和功能之间作一折衷，使小应用程序也能有希望地访问文件、硬盘和数据库。现在，JavaSoft 公司已经开始了这方面的工作。

目前，可用于 Java Beans 开发的工具有：Java Soft 的 BDK (Beans Development Kit)，IBM 公司的 VA (VisualAge for Java)。另外，在 JDK1.1 及其后续版本中，有一个引人注目的类库就是 java.beans。

第三，JavaSoft 公司在 Java 开发工具的发展方面是不遗余力的。他们主要提供基于各种操作系统平台的 Java 开发工具包 JDK (Java Development Kit)。从最初的 JDK1.0，到 JDK1.0.2，到 JDK1.1 (beta)，而后是 JDK1.1 的正式版。到了六月份，JDK 已经发展到了 JDK1.1.1 和 JDK1.1.2。并且，JDK1.1.1 还推出了日文版。只有短短的几个月的时间，随着版本的不断更新，JDK 中的类库越来越大，功能也越来越完善。JDK1.0 只包含了最基本的类库，如：java.lang (定义 Java 语言)，java.applet (用于开发 Java 小应用程序的类库)，java.awt (Java 提供的用于开发 GUI 程序的类库，AWT, Abstract Window Toolkit)，java.net, java.io, java.util 等等。到了 JDK1.1，在 JDK 的类库中又增加了以下的几个类库：如 java.sql (提供访问 SQL 数据库的机制)，java.beans, java.rmi (remote method invocation 远程方法调用)。并且，在 JDK1.1 中，对 AWT 的事件处理机制作了较大的改进，其事件处理已不再象 JDK1.0.2 那样复杂，并且其形式和机制也更接近于 MOTIF 的事件处理方式或 Windows 的消息传播机制。这使人们能更加容易地编写 Java 的 AWT 程序。Java Soft 公司在开发 JDK 的同时，也推出利用 JDK 来开发应用程序，典型的如完全是用 Java 语言来编写的浏览器，Hot Java 1.0 (beta)，它就使用了不少 JDK1.1 中的类。

Symantec 公司的 Cafe Lite 是一个 32 位的 Java 程序的集成开发环境，能方便地建立工程文件，能方便地查看源文件和特定的类定义，支持 32 位的源码调试，用户界面友好。而 Microsoft 并不想自己被排除在 Java 俱乐部之外，他利用自己在可视化开发平台方面的优势，很快推出了 VJ++1.0，把 Java 纳入了 Visual 的系列产品之中。VJ++ 目前最新的版本是 VJ++1.1。同时，Microsoft 又不断加强 ActiveX 对 Java Beans 的攻势。在这里，我们比较一下 Java Soft 和 Microsoft 目前在 Java 和 Internet 上的竞争，也许是很有意思的：

	Microsoft	Java Soft
平台	Windows, Windows NT	各种平台
开发工具	Visual J++	WorkShop
浏览器	Internet Explorer	Hot Java
软构件	Active X	Java Beans

第四，基于 Java 的应用的开发，其出发点是构造更好的网络使用环境。它包括两个方面的内容。一是 Java 如何利用现有的极其庞大的计算机资源，如 Java 与本机现有软件进行互操作的问题；Java 通过网络访问数据库的问题；另一是如何面对新的应用领域，如如何实现网络多媒体；如何与 VR (虚拟现实) 技术相结合；在 Web 上实现 VR；Java 与 CORBA (网络中间件) 的集成，实现分布式 Java 编程等等。基于以上的认识，我们可以发现，目前基于 Java 的应用程序的开发主要集中在四个方向。一是对网络多媒体应用的开发 (如 Media play, animation 和 frame)；二是和 VR 技术的结合；三是提供对企业网络应用和管理的解决方案；四是网络数据传输和访问。

如何把 Java 技术应用于多媒体的研究和开发，是 Java 在应用领域的一大发展方向。现在已经有一些公司 (主要是 SUN, SGI 和 Intel 公司) 开始在这一方面先行一步，合作制定了 Media Framework 的标准。如 Intel 公司，提供了几个用 Java 编写的可以利用客户机的浏览器来浏览多媒体的工具包，Media Framework Package, Intel Animation for Java Package, Intel Spatial Audio for Java Package 等等。Java Soft 公司也提供了有关多媒体播放和通信的 API 和技术文档。这些 API 都允许开发者和用户方便灵活地利用 Web 上范围广泛的、丰富的、交互式媒体资源，如媒体框架、动画、2D、3D，通过 Web 实现通话、演讲和协作等等。可以预计，随着网络多媒体技术的迅速发展，计算机正在进入传统的电视机市场。

同时也有其他的一些公司在展开对电视机顶盒的研制,其目的是使电视机能访问 Internet。这两者的共同发展,最终将达到计算机和电视机的一体化。

另外,象 SGI 公司利用自己在图象处理和 VR 技术方面的一贯优势,在把 Java 和 VRML 的结合方面进行了一定的努力,最新的 VRML2.0 版,就提供了 VRML 与 Java 和 Java SCRIPT 的接口。在浏览器方面,Netscape 和 Internet Explore 都对 Java 进行了不遗余力的支持。

Java Soft 公司为了在公司和企业中推广对 Java 技术和产品的应用,推出了许多面向企业应用的 Java API。如 JNDI(Java Naming and Directory Interface),提供一致的、符合工业标准的、具有从 Java 平台到商业信息的无缝连接功能。这样,JNDI 允许开发者在具有多个命名和多个目录的服务器企业网上传送应用程序时,能有唯一的访问路径。又如 Java 的面向分布式编程的 IDL(Interface Definition Language)系统 Alpha 2.0,提供了一种利用工业标准 IDL 透明连接 Java 客户和网络服务器的方式。还有 JDBC(Java Database Access API 1.10 版)为程序员提供了访问各种相关数据库的单一界面,以及 JRMI(Java 远程方法调用)和 Java 对象串行化(Object Serialization)等等。

另外,为数众多的公司和开发小组也提供了大量的利用 Java 来传输和访问数据的软件包。如 InfoSpace 公司用 Java 技术开发了通过 Web 浏览器访问数据库的 SpaceSQL;Open Connect Systems 公司提供的基于 Web 浏览器的企业网基本结构;Web-Connect。还有如 Andersen Consulting 的 via World Network 是一个完全用 Java 来开发的用于旅游的计划 and 预定系统,它揭示了 Java 在诸如购物、订票等方面应用的广阔前景。并且,由于这些应用往往需要诸如网络数据库访问、网络多媒体、虚拟现实等方面的技术。因此,这些应用的发展也将带动此类技术的发展。

所有这些,都充分表明了一个十分明确的发展方向:使 Java 成为统一的网络平台。它的内容可能包括 Java OS;完全用 Java 语言开发的浏览器(如 HotJava);或是对 Java 完全支持的浏览器(如 Netscape,Internet Explorer 等等);网络上大量的可以随时下载的 Java 应用程序;用 Java 开发类似 CORBA 的中间件;还有就是专门运行 Java 程序的计算机如 NetPC 的 NC。在这个 Java 网络平台上,将真正实现整个计算机网络上的数据共享和服务共

享。

三、挑战和机遇

Java 的迅猛发展,既给我们带来了挑战,又为我们带来了机遇。Java 促使 Internet 使用范围的迅速扩大,日渐进入寻常百姓人家。同时,由于网络上的信息千变万化,各种软件将以比原来更快的速度更新,因此 Java 促使我们考虑改变原有的软件设计和开发模式,考虑如何用新的目光、新的思想来适应新的软件大市场。

首先,Java 是新事物,它的发展历史并不长,国际上对 Java 的研究和应用也是处于刚刚起步的阶段。并且,Java 所解决的和它正在加以解决的问题对于我们来说并不太陌生;象平台无关性问题,计算机网络上的互操作问题等等,在 Java 出现之前,也都已经开始研究了。如果我们现在就迎头追赶的话,我们和国际先进水平的差距并不很大。并且,对 Java 的开发和研究,将有助于缩短我国在网络软件开发方面与国外的差距,并在庞大的网络软件市场上占据一席之地。

其次,国内网络应用软件的开发还很薄弱,正象是在一张白纸上画画,不象许多国外的大公司那样,需要更多地考虑如何把 Java 与本公司现有的产品有机地结合起来。而我们则可以尽情地发挥我们的想象力和创造力,用 Java 为我国的计算机网络提供更多、更美好的应用软件,扩大我国计算机网络的应用范围和层次。并且,由于 Java 是与平台无关的,这对于象我国那样硬件基础比较薄弱的国家,开展网络软件的研制和开发是一个很有利的条件。目前,我国已经建立了 ChinaNet,但是,由于缺乏网络应用软件,网络的应用还远未达到它应有的水平。因此说,在我国开展对 Java 的研究,有很大的应用需求。

第三,人们常说,Java 就是没有指针的 C++,这话虽不准确,但也有一定的道理。确实,Java 语言的语法规则很象 C++,Java 程序初看起来也很难与 C++ 程序分清。Java 语言成功地借鉴了 C++ 许多优点,同时对 C++ 所存在的不足,又精心加以防止。由于 Java 的类 C 风格,使 C 和 C++ 程序员很容易入门。我国目前有大量的熟知 C、C++ 的专业人员,对他们来说,掌握 Java 语言是一件很容易的事。因此,可以说我国已拥有相当数量的适于 Java 研究和开发的人员。

还有一个有利的条件是,目前,国内的许多单位已经或正在建立本单位的计算机网络,并且大都能

数据库

知识发现

KDD

机器学习

②

计算机科学 1997 Vol. 21 No. 6

KDD: 数据库中的知识发现

KDD: Knowledge Discovery in Database

朱廷勋 高文

(先进人机通信技术联合实验室 北京 100080)

S-9

TP311.13

摘要 KDD(Knowledge Discovery in Database) is a new research field which comes into being with the development of Database and Artificial Intelligence, it extracts useful information from large Database. In this article, we will introduce the basic concept and process of KDD and the differences among KDD, Machine Learning and Database report. At last, we will give some conclusion of KDD.

关键词 Knowledge Discovery in Database, Data Mining, Database, Machine Learning

一 引言

随着数据库技术的不断发展及数据库管理系统的广泛应用,数据库中存储的数据量急剧增大,但目前数据库系统所能做到的只是对数据库中已有的数据进行存取,人们通过这些数据所获得的信息量仅是整个数据库所包含的信息量的一部分,因为目前用于对这些数据进行分析处理的工具却很少,又有局限性。然而,隐藏在这些数据之后的更重要的信息是关于这些数据的整体特征的描述及对其发展趋势的预测,这些信息在决策生成的过程中具有重要的参考价值。

在数据库技术飞速发展的同时,人工智能领域的一个分支——机器学习的研究也取得很大进展。自 50 年代开始机器学习的研究以来,在不同时期的

研究途径和目的也不尽相同,一般大致可分为三个阶段,其研究内容分别为:神经模型和决策理论、概念符号获取及知识加强和论域专用学习。根据人类学习的不同模式人们提出了很多机器学习方法,如:实例学习、观察和发现学习、神经网络和遗传算法等等。其中某些常用且较成熟的算法已被人们运用于实际的应用系统及智能计算机的设计和实现中。

正是由于数据库技术和机器学习技术的发展,也是为了满足人们实际工作中的需要,数据库中的知识发现(KDD)技术逐渐发展起来。KDD 也有人称之为数据挖掘(Data Mining),实际两者是有区别的,但一般可以不加区别地使用,在本文我们统一以 KDD 称之。

二 KDD 定义

从开始到现在,人们给 KDD 下过很多定义。随

够通过 ChinaNet 与 Internet 相连。这使我们拥有一个良好的 Java 环境。同时,由于国际上基于 Java 的应用尚处于试验阶段,在 Internet 上有大量的有关 Java 的免费软件可以下载。这一方面能使我们了解到国外的最新动态,对我们开展对 Java 的应用研究有借鉴作用;另一方面也使我们能有效地利用别人的成果,既少走弯路,又能在一个较高的水平上开始对 Java 进行研究和应用。

当然,开展对 Java 的研究和应用,应该注意的一个问题是,切忌各行其是,大家都在低水平上重复。应该有规划、有目标、有要求;要有组织、有分工、有协作,发挥集团作战的优势,共同把 Java 的应用

和研究搞上去,为我国的网络应用的推广,走出一条捷径。

参考资料

- [1] "Moving to JDK 1.1: Using the delegation event model to create custom AWT components", Merlin Hughes, Java World, 1997-05
- [2] "Graphic Java, Mastering the AWT", David M. Gerry, Alan L. McClellan, Sunsoft Press, 1997
- [3] "JDK1.1 Documentation", Doug Kramer, Sun Microsystems, Inc.
- [4] "Web Site Programming with Java", David Harms et al., McGraw-Hill, 1996
- [5] "Java Beans White Paper", Sun Press, 1996