About Worldwide Software

Read topics from the **WorldWide Software Overview** if you wish to create applications that are adapted to regions other than your own or that work with any non-Roman script system. The information is essential if you would like to plan and provide for such development in future versions of your applications. Furthermore, you may want to take advantage of several other text-handling capabilities described in the **WorldWide Software Overview** and not available elsewhere in the Macintosh Tool Box.

The **WorldWide Software Overview** provides the essential background to developing Macintosh software for world-wide markets. It introduces you to the use of the Macintosh worldwide software, whose specific features are provided by the **Script Manager**, the **International Utilities Package**, the international and keyboard resources, and script systems. You can use worldwide software to make your applications compatible in the global market.

The **WorldWide Software Overview** provides an overview of the **Script Manager** routines and data structures that let you represent scripts on the Macintosh. (**Scripts,** in this context, are writing systems such as Roman, Japanese, and Arabic that are used to represent human languages.) The **Script Manager** allows you to deal with script-related issues such as character representation, text direction, contextual forms, diacritical marks, uppercase and lowercase characters, character reordering, word demarcation, and text alignment.

The **WorldWide Software Overview** gives an overview of the **International Utilities Package** routines and data structures that permit you to sort strings and to format dates, times, currency, and numbers according to the conventions of the script, language, or region of the software market you address. Throughout this section, **language** refers to the whole body of written words and of methods of combining words used by a particular group of people, and **region** denotes a linguistic or cultural entity that may or may not correspond to a geographic area.

The **WorldWide Software Overview** gives details on the international and keyboard resources that help you specify information that pertains to a particular script, language, or region. Such information includes fonts, long and short date formats, keyboard layouts, preferred sorting order, and relation-ships between scripts, languages, and regions. Other pertinent data is delineated in resource tables for character type, case conversion, and word breaks. The resources also allow you to specify tokens, character set encodings, and keyboard mapping information that includes hardware-specific modifications to keyboard layouts.

The **WorldWide Software Overview** introduces the Macintosh **script systems**-collections of software facilities that work with the **Script Manager** to provide for basic differences between writing systems, such as character sets, fonts, keyboards, text collation, and word breaks. Examples of script systems are Roman, Japanese, Arabic, Traditional Chinese, Simplified Chinese, Hebrew, Cyrillic, Thai, and Korean.

Finally, The **WorldWide Software Overview** furnishes an extensive set of guidelines for adapting your applications to other languages and regions and

THINK Reference © 1991-1992 Symantec Corporation

for writing software for other scripts.

The information in the **WorldWide Software Overview**, particularly, on the Macintosh Script Management System and the concepts underlying the Macintosh script systems, provides you with the understanding you need to make your applications run on Macintosh computers with multiple script systems installed. To make your applications work in a region other than your own, you will find the **International Utilities Package**, the international resources, the keyboard resources, and the concepts underlying the localization process especially useful.

The essential technical reference on the features in worldwide software prior to system software version 7.0 is the beta draft of *Macintosh Worldwide Development: Guide to System Software*. That manual, currently available through APDA, supplements the information here. It covers in depth the **Script Manager** routines and data structures and the international and keyboard resources. Together with this overview, *Macintosh Worldwide Development: Guide to System Software* tells you what you need to know about making your software compatible worldwide.

You should also be familiar with the **User Interface Guidelines** and with *Human Interface Guidelines: The Apple Desktop Interface,* available through Addison-Wesley Publishing Company, Inc., and with the following information:

- text manipulation functions in **QuickDraw**
- the **Font Manager**'s support for **QuickDraw**
- the **Binary-Decimal Conversion Package**

The **TextEdit** section provides some examples of how to use the **Script Manager**.

The Macintosh **worldwide system software** helps you address the issues you'll encounter when you design your applications to be compatible with regional, linguistic, and script differences around the globe. It enables you to create applications that run in other regions or work with different scripts. Worldwide system software consists of the Macintosh Script Management System (the **Script Manager** and one or more script systems) and the **International Utilities Package**, international resources, keyboard resources, and certain keyboard-handling routines, etc.

As you enter the process of developing applications for worldwide markets, it is important to consider variations that are specific to script, language, and region. Scripts may differ in the direction in which their characters and lines run, the size of the character set used to represent the script, and context sensitivity. Examples of script-specific features include text display, text rendering, text editing, fonts, input methods, and character set encoding. Examples of language-specific features include sorting order and word boundaries. Region-specific features include date and time format, number format, and case conversion.

> **Note:** Read **An Introduction to Scripts** to acquaint yourself with the concepts you must understand if you want to create software for writing systems other than your own.

THINK Reference © 1991-1992 Symantec Corporation

The Graphic Representation of Languages

The worldwide system software-especially the Macintosh Script Management System-deals primarily with the graphic representation of language, not with spoken language. This has implications for the treatment of languages, including the numeric codes assigned to represent each language. A spoken language that may be written in more than one script is treated on the Macintosh as several languages: one for each script in which the language is written. In some cases, this distinction is already present in the names of spoken languages. For example, Romanian and Moldavian are essentially the same spoken language; however, in Romania this language is written in Roman script, whereas in the adjacent Soviet province of Moldavia, this language is written in Cyrillic script. In other cases, this distinction is not present: the official language of Malaysia may be written in either Roman or Arabic script, but the spoken language is called Malay in either case. The Macintosh Script Management System distinguishes the written versions with **language codes** (numbers used to indicate particular languages on the Macintosh) such as langMalayRoman and langMalayArabic.

Localized Versions of the Macintosh System Software

Localization is the process of adapting software to a particular region and language. This can include translating text to another language, using the region's date, time, and number formats, adapting icons and other graphic elements to the cultural conventions of the target region, and so on. Although localization often requires language capability in the system software and may require script capability, it is not synonymous with adding script or language capability to a system. Localization is an operation that can potentially affect every part of the system software. For example, the Japanese Script System might be localized for France so that the text in the Japanese Script System control panels would appear in French.

The Macintosh Script Management System accommodates worldwide differences with the concept of regions, which provide a finer and more complex level of granularity than script and language. For example, the French language is used in France, in parts of Belgium, Switzerland, and Canada, and in other countries such as Luxembourg, Haiti, Mali, Zaïre, Tahiti, and Vanuatu. Each of these areas may have different conventions for time, date, and number formats. Some differences may also occur in the behavior of the written language. For example, in France, accents on most characters are generally omitted if the character is written in uppercase; in Quebec, the accents are usually preserved. A **region code** is a number that may designate a region that is conceptually smaller or larger than a country (for example, "French Swiss" or "Arabic"), as long as the region shares the characteristics described above.

Localized versions of the Macintosh system software (including the Roman Script System, the Macintosh Operating System, the Toolbox, and so forth) combine information specific to scripts, languages, and regions.

> **Note:** Currently the following localized versions of the Macintosh system software are available: Arabic, Australian, British, Croatian, Danish, Dutch, Farsi, Finnish, French, French Belgian and Luxembourgian, French Canadian, French Swiss, German, German Swiss, Greek, Hebrew, Hindi (India), Icelandic, Irish, Italian,

THINK Reference © 1991-1992 Symantec Corporation

Japanese, Korean, Maltese, Norwegian, Portuguese, Simplified Chinese, Spanish, Swedish, Thai, Traditional Chinese, Turkish, and U.S.

For more on the process of localization in worldwide software, see the beta draft of *Guide to Software Localization,* currently available through APDA.

Multiple Script Systems and Multiple Languages

At least two script systems are always present when a non-Roman script system is installed. For example, the Japanese system software is the combination of the U.S. system software (which includes the Roman Script System, the Macintosh Operating System, the Toolbox, and so forth) and KanjiTalk (the Japanese Script System), all of which are localized for Japan. Localized versions of the Macintosh system software with more than one script system installed-such as the Japanese system-have been adapted to particular regions, languages, or countries. However, systems such as the French and Turkish versions are simply localized variations of the U.S. system software, which do not include a second script system. With system software version 7.0 , script systems may be installed either as a **secondary script** (also called auxiliary script), which just provides script capability, or as the primary script (also called **system script**), which affects system defaults and is the script used for dialog boxes, menus, and alerts.

The **Script Manager** has always supported the simultaneous use of more than one non-Roman script. System software version 7.0 makes it easier for users to install multiple script systems, and in version 7.0, script systems function properly even if they are not the system script. These enhancements provide increased opportunity for your applications to take advantage of the **Script Manager**'s handling of multiple scripts.

You may decide to support multiple languages in your applications. Word processors, for example, might tag a text run with a language attribute in a manner similar to style attributes; this language tag could then govern the behavior of spelling checkers, hyphenators, and so on. In addition, you can let users choose a presentation language for your application, that is, the language for menus, dialog boxes, and alerts.

Identifying Scripts, Languages, and Regions

Scripts, languages, and regions are organized into a strict hierarchy. The Macintosh Script Management System distinguishes languages at a finer level than usual. If a language in the traditional sense can be written in more than one script, each language and script combination is called a separate language in this context.

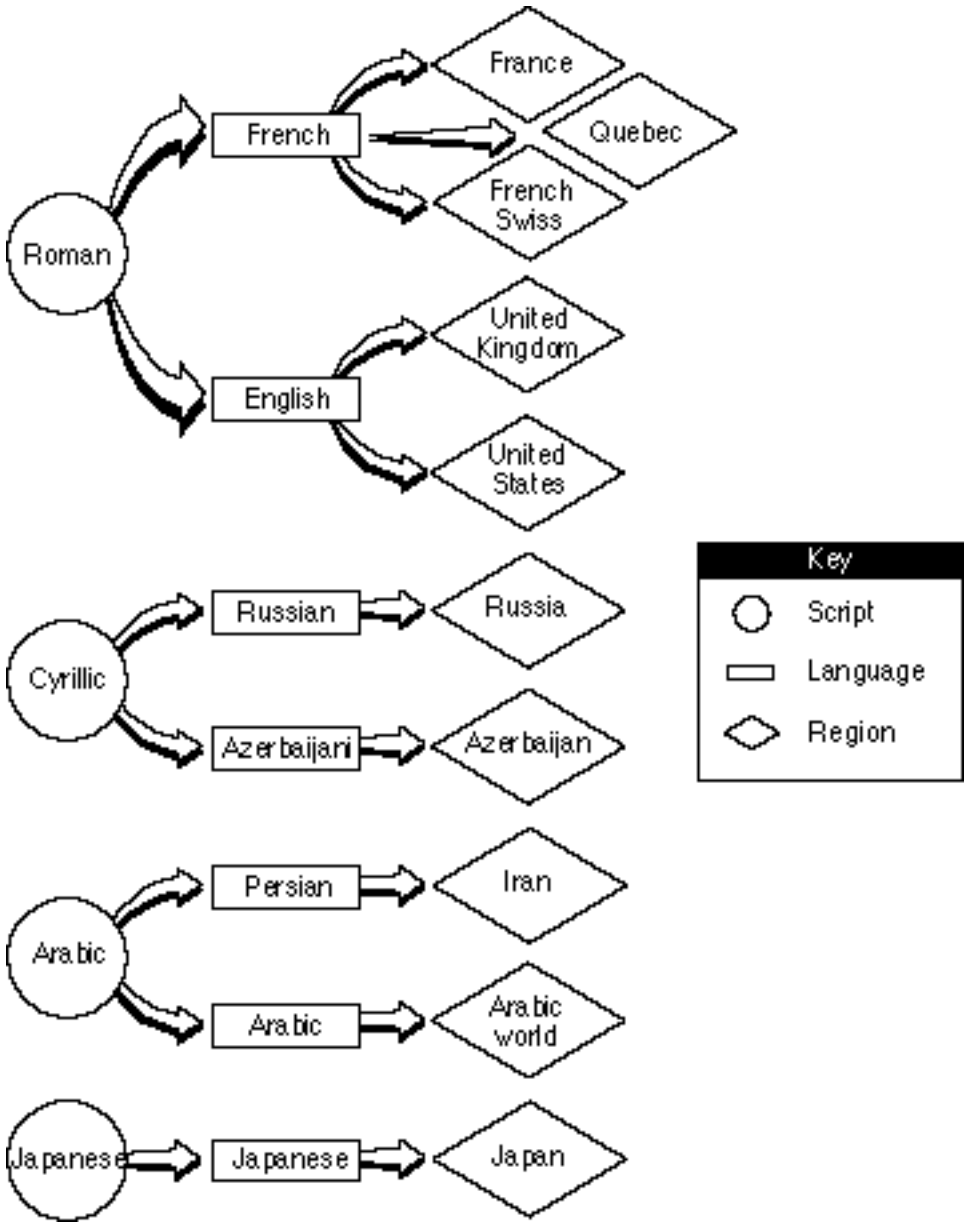
Three basic principles underlie the hierarchy of script, language, and region:

- Every different character set encoding has a different **script code.** (For historical reasons, this is not strictly true for some localized versions of the Roman Script System.)
- Languages always belong to a particular script.
- Region codes always belong to a particular language (a localized version of the system software is for a particular language in a particular country or other region). Several regions may be associated

THINK Reference © 1991-1992 Symantec Corporation

with a particular language.

The Figure below illustrates the script, language, and region hierarchy.



The script, language, and region hierarchy

With System 7.0 several changes and additions have been made to the script, language, and region codes used by the Script Management System. Most of the changes are backward-compatible. See the **Script Manager Data** for a complete list of these codes, including changes to the constant code names for scripts, languages, and versions.

THINK Reference © 1991-1992 Symantec Corporation