Paper 336-2010

Best Practices for a Non-English BI Deployment

Sunny Li and Holmes Niu, SAS Institute Inc., Beijing, China

ABSTRACT

SAS products are used by customers around the world. This paper describes how to deploy and configure business intelligence products for non-English environments (such as Japanese, Chinese, French, etc.) and how to configure SAS Enterprise BI suite so that it can support multiple languages simultaneously. This paper will provide the guidelines and best practices of installing and configuring SAS BI system under multiple language environments. Highlights include: how to install and configure a SAS BI system in a specified language environment; how to add a new SAS Application server to support additional languages; how to deploy a Unicode SAS BI system to access and explore multilingual data with SAS BI client.

INTRODUCTION

The SAS 9.2 System solves business problems for customers around the world. SAS software adapts to the language and cultural preferences of its customers. This article focuses on the installation and deployment of the SAS 9.2 BI platform under non English language environments.

DEPLOY THE SAS BI PLATFORM IN A SPECIFIED LANGUAGE ENVIRONMENT

The SAS® Foundation and other SAS products have been translated and made available for many locales so that people can work with SAS in their own native language environment. These locales include simplified Chinese, traditional Chinese, Japanese, Korean, French, German, Italian, Spanish (Castilian), Russian, Danish, Swedish, Hungarian, and Polish.

When deploying a SAS BI platform, the SAS® Deployment Wizard configures the SAS language settings according to the operating system’s language setting automatically. But in some cases, to deploy a non-localized version of the SAS System using the SAS Deployment Wizard, a customer might need to manually change the language settings after a deployment. This article will go over the following four scenarios which the customer could possibly encounter:

1. Scenario I: Deploy a localized version of the SAS System by the SAS Deployment Wizard.
2. Scenario II: Deploy a non-localized version of the SAS System by the SAS Deployment Wizard.
3. Scenario III: Change the default encoding selected by the SAS Deployment Wizard during deployment.
4. Scenario IV: Change the default locale selected by the SAS Deployment Wizard during deployment.

SCENARIO I: DEPLOY A LOCALIZED VERSION OF THE SAS SYSTEM BY THE SAS DEPLOYMENT WIZARD

The easiest way to deploy a localized version of a SAS System is to use the default language setting in the SAS Deployment Wizard. There are still some steps that should be performed before and during the deployment.

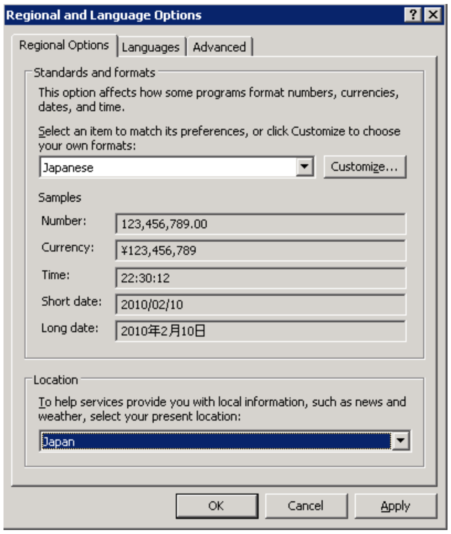
1. Configure the operating system's language setting.

The SAS Deployment Wizard assumes that the operating system (OS) regional and language settings match the user's preference for the SAS deployment.   If, for some reasons, the target operating system has a different language setting from the expected language for SAS, then you must switch the operating system's language settings according to the expected SAS BI language environment. There are differences between the configuration and language settings of Windows and UNIX operating systems.

- **Windows**

Language settings in different version of Windows can be different and beyond the scope of this article. We only take Windows MUI 2003 server for example to demonstrate how to check and change the language and regional settings. The following steps show how to set Japanese (the locale is ja_JP) as the Windows system language.

1) Click **Regional and Language Options** from the Control Panel. In the **Regional Options** tab, you can review which locale is currently used (Figure 1).



Regional and Language Options

Standards and formats

This option affects how some programs format numbers, currencies, dates, and time.

Select an item to match its preferences, or click Customize to choose your own formats:

Japanese

Samples

Number: 123,456,789.00

Currency: ¥123,456,789

Time: 22:30:12

Short date: 2010/02/10

Long date: 2010年2月10日

Location

To help services provide you with local information, such as news and weather, select your present location:

Japan

Figure 1. Regional and Language Options

2) In the **Regional Options** tab, select *Japanese* in the **Standards and formats** section. Select *Japan* as the **Location**.

3) In the **Languages** tab, do not check **Install files for complex script and right-to-left languages** but instead check **Install files for East Asian languages**. Select *Japanese* as Language used in menus and dialog boxes.

4) In the **Advanced** tab, select *japanese_Japan* in the **Language for non-Unicode programs** section and check the box in **Default user account settings**.

- **UNIX**

Setting the locale is similar on most UNIX systems. This paper uses HP-UX as an example to demonstrate how to change your system locale on UNIX. The executable *locale* shows your current locale:

```
$ locale
```

The following command displays all the locales that are currently installed on your machine:

```
$ locale -a
```

The following command sets the current locale:

```
$ export LANG=locale
```

The general syntax for locale values is: language [_region][.charset]. For example, in ja_JP.SJIS ja_JP denotes the language and country and SJIS sets the encoding.

The table below illustrates how to set this variable under different shell environments.

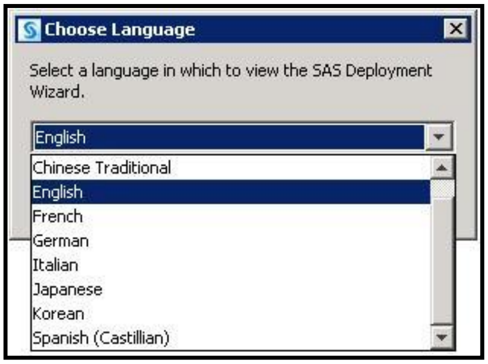
UNIX and Linux	

Bourne Shell	`$ LANG=ja_JP.SJIS`
	`$ export LANG`

C Shell	`$ setenv LANG=ja_JP.SJIS`

2. **Choose the language for the SAS Deployment Wizard user interface (UI)**.

The first step is to choose a language for the installation. The SAS Deployment Wizard supports nine languages for the installation as Figure 2 shows. If the language of the OS is available in the list of languages for the SAS Deployment Wizard, then choose the language that is compatible with the OS' language setting in the drop list. If the OS' language is not in the list, then choose English. For example, Polish is a language that is supported by SAS Foundation, but it is not available as an installation language for the SAS Deployment Wizard. Because the language selection will be also used for the BI configuration, it is recommended that you choose the language that is compatible with the OS' language setting or English.



Choose Language

Select a language in which to view the SAS Deployment Wizard.

English

Chinese Traditional
English
French
German
Italian
Japanese
Korean
Spanish (Castillian)

Figure 2. Choose a Language to View in the SAS Deployment Wizard UI

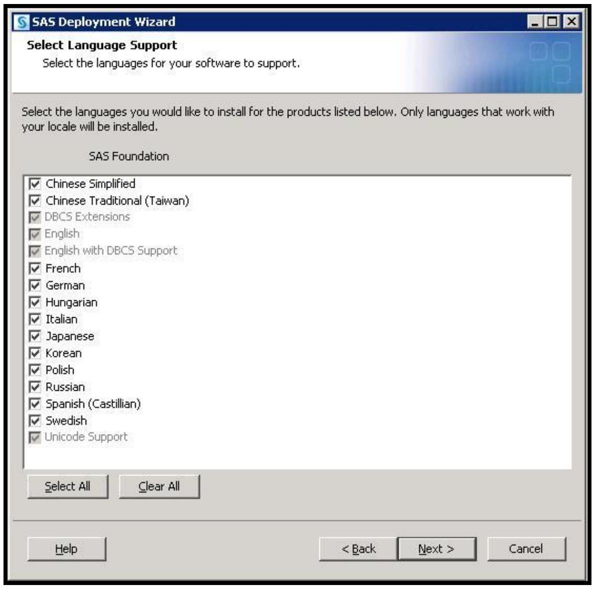
3. Select the supported languages to install.

The SAS Deployment Wizard allows a customer to install several language versions from the Select Language Support dialog box (Figure 3).

For Windows, the SAS Deployment Wizard selects the languages to install that are compatible with the operating system's language setting from the checked languages in this dialog box. For example, if all the languages are selected in this dialog box (Figure 3) and the OS locale is fr_FR (French), then the SAS Deployment Wizard installs French, German, Spanish, and Italian because those four languages are compatible with the OS' language setting.

For UNIX, all selected languages in the Select Language Support dialog box are fully installed.

DBCS Extensions, English, English with DBCS Support, and Unicode Support are checked and grayed out in the dialog box because the SAS Deployment Wizard will install DBCS and Unicode support with the English product by default.



SAS Deployment Wizard

Select Language Support

Select the languages for your software to support.

Select the languages you would like to install for the products listed below. Only languages that work with your locale will be installed.

SAS Foundation

☑ Chinese Simplified
☑ Chinese Traditional (Taiwan)
☑ DBCS Extensions
☑ English
☑ English with DBCS Support
☑ French
☑ German
☑ Hungarian
☑ Italian
☑ Japanese
☑ Korean
☑ Polish
☑ Russian
☑ Spanish (Castillian)
☑ Swedish
☑ Unicode Support

Figure 3. Select Language Support for SAS Foundation

The SAS Deployment Wizard will create the folders for the selected languages under the directory named *<SASRoot>\nls\*. Note that *SASROOT* is the location of your installed SAS Foundation directory. For example, in Windows, 'c:\program files\SAS\SAS Foundation\9.2' is the default location of SASROOT.

The default language version for a SAS session is set by the SAS Deployment Wizard according to the OS' language setting.

SCENARIO II: DEPLOY A NON-LOCALIZED VERSION OF THE SAS SYSTEM BY THE SAS DEPLOYMENT WIZARD

In a non-localized version of the SAS System, English is the default language of the UI, message files, Output Delivery System (ODS) templates, and so on. However, data in those languages can be correctly processed if the SAS session supports the locale that is associated with the language. The SAS Deployment Wizard doesn't create a folder for a non-localized language in the *<SASRoot>\nls\* directory.

For Windows, the SAS Deployment Wizard modifies the English configuration file named sasv9.cfg in the folder named *<SASRoot>\nls\en* to support the non-localized language. An option "Locale" is added into the English configuration file to specify the locale value for that non-localized language.

For UNIX, a customer must manually add the "Locale" option in the file named *<SASROOT>\ sasv9_local.cfg* to support the non-localized language. For example, Greek is not localized for the SAS System when SAS BI is deployed in a Greek environment. The option "-locale el_GR" should be added in the corresponding configuration files so that the SAS session can process the Greek data.

Note: If a SAS session does not support the locale that is associated with the non-localized language (for example, Azerbaijani), then the default locale, en_US, and encoding, WLatin1, will be used. Data of such non-supported locales cannot be processed in a SAS session.

For details about locales that are supported by SAS 9.2, see the sections about the LOCALE= system option in the *SAS® 9.2 National Language Support (NLS): Reference Guide*.

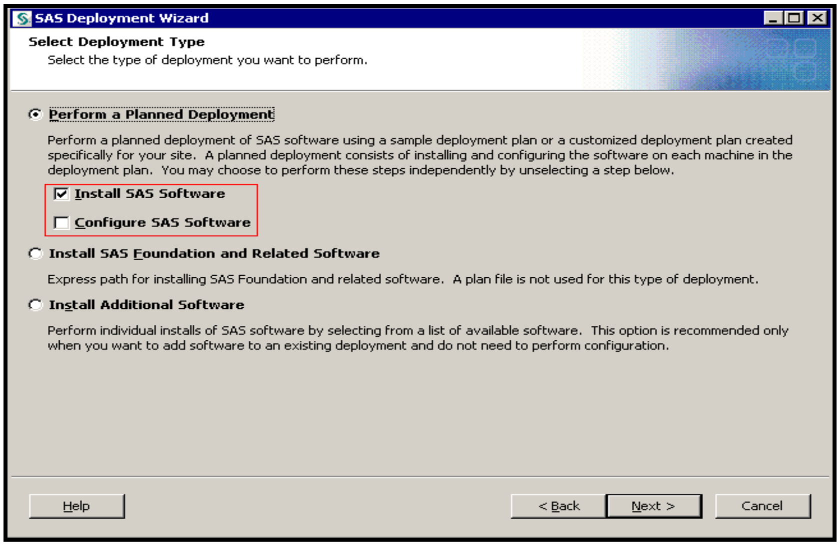
SCENARIO III: CHANGE THE DEFAULT ENCODING SELECTED BY THE SAS DEPLOYMENT WIZARD DURING DEPLOYMENT

The SAS Deployment Wizard specifies the "Locale" option[1] in the SAS configuration file according to the OS' language settings during deployment. If the encoding is not specified explicitly in the configuration file, then the locale value determines the encoding under which the SAS session runs. Typically, a SAS session uses one of many encodings that are compatible with the locale as the default encoding. For example, latin9 and latin1 are both compatible with the locale de_DE (German_Germany locale). SAS uses latin9 encoding in UNIX as default encoding

[1] The locale option is indeed not set in the configuration file for Asian languages. However, it shows up in the log when you run a PROC OPTIONS. DBCSTYPE and DBSLANG determine the locale.

for locale de_DE. Sometimes, customers might need to explicitly specify an encoding for a SAS session to run instead of using the default encoding that is implied by the “Locale” option. In this case, the “encoding” option should be added into the SAS configuration file during deployment using the following steps.

1) Launch the SAS Deployment Wizard to install the SAS software. Only check **Install SAS Software** in the Select Deployment Type dialog box (Figure 4).



SAS Deployment Wizard

Select Deployment Type

Select the type of deployment you want to perform.

● **Perform a Planned Deployment**

Perform a planned deployment of SAS software using a sample deployment plan or a customized deployment plan created specifically for your site. A planned deployment consists of installing and configuring the software on each machine in the deployment plan. You may choose to perform these steps independently by unselecting a step below.

☑ **Install SAS Software**

□ **Configure SAS Software**

○ **Install SAS Foundation and Related Software**

Express path for installing SAS Foundation and related software. A plan file is not used for this type of deployment.

○ **Install Additional Software**

Perform individual installs of SAS software by selecting from a list of available software. This option is recommended only when you want to add software to an existing deployment and do not need to perform configuration.

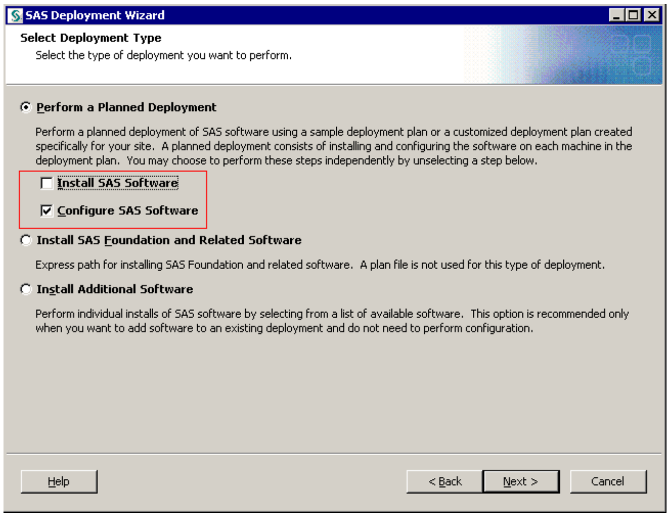
Figure 4. Select Deployment Type. Check only **Install SAS Software.**

2) Add an “encoding” option in the SAS configuration file after the installation.

After the installation, add the required “encoding” option into the configuration file. For the above example, add the following line to *<SASROOT>/ sasv9_local.cfg* (for UNIX) or *<SASROOT>/nls/de/sasv9.cfg* (for Windows).

```
-encoding latin1
```

3) Rerun the SAS Deployment Wizard to configure the SAS software. Only check **Configure SAS Software**. Do not check **Install SAS Software** in the Select Deployment Type dialog box (Figure 5).



SAS Deployment Wizard

Select Deployment Type

Select the type of deployment you want to perform.

● Perform a Planned Deployment

Perform a planned deployment of SAS software using a sample deployment plan or a customized deployment plan created specifically for your site. A planned deployment consists of installing and configuring the software on each machine in the deployment plan. You may choose to perform these steps independently by unselecting a step below.

□ Install SAS Software

☑ Configure SAS Software

○ Install SAS Foundation and Related Software

Express path for installing SAS Foundation and related software. A plan file is not used for this type of deployment.

○ Install Additional Software

Perform individual installs of SAS software by selecting from a list of available software. This option is recommended only when you want to add software to an existing deployment and do not need to perform configuration.

Figure 5. Select Deployment Type. Check only **Configure SAS Software.**

For information about session encodings that are supported by SAS 9.2, see the section named “Encoding Values for a SAS Session” in the *SAS® 9.2 National Language Support (NLS): Reference Guide.*

SCENARIO IV: CHANGE THE DEFAULT LOCALE SELECTED BY THE SAS DEPLOYMENT WIZARD DURING DEPLOYMENT

There are some cases in which the locale you want to use is not the default locale that is selected by the SAS Deployment Wizard. For example, in Windows, if you want to use the locale fr_CA, then you set the OS locale to fr_CA. After installing the SAS BI platform, the SAS Deployment Wizard will use the fr_FR locale instead of the fr_CA as the SAS session’s locale. In order to use the fr_CA locale, the sasv9.cfg configuration file needs be changed manually. The installation and configuration steps are similar to those in scenario III.

1) Invoke the SAS Deployment Wizard to install the SAS BI platform, not to configure it.

2) After the BI platform has been installed successfully, open sasv9.cfg under <*SASROOT*>*\nls\fr.* Replace `–locale French` with `–locale fr_CA`.

3) Invoke the SAS Deployment Wizard again to configure the SAS BI platform.

ADD A NEW SAS APPLICATION SERVER TO SUPPORT AN ADDITIONAL LANGUAGE FOR AN EXISTING BI PLATFORM

The sections above described how customers install and configure the BI platform for one language. There will be cases where a customer wants to use the BI platform that is set up for one language to process data in another language. Consider this business scenario. An international company is headquartered in Korea and has installed a SAS BI platform with the korean_KOREA locale on Windows. The default encoding that is used by the SAS BI platform is euc-kr. If the company opens a new branch in China, then they will have new requirements. These requirements are as follows:

1) The same BI platform should be able to process Chinese and Korean data.
2) Korean users can view the reports containing Korean data and Chinese users can view the reports containing Chinese data.

To meet the requirements, it is recommended that you add a second SAS application server to process the Chinese data. All stored processes, jobs, and cubes that are used to process the Chinese data should run on the second application server. The original SAS application server would service Korean users and the new application server would service Chinese users.

We take the above business scenario as an example to describe the steps. Suppose the name of the original application server is SASApp_KO and the name of the additional application server is SASApp_ZH. The two application servers run on the same machine as where the SAS® Metadata Server is running. They use the same object spawner.

Add the additional SAS application server named SASApp_ZH. The steps about how to add new SAS application server can be found in Chapter 10, "Managing SAS Application Servers" in the *SAS® 9.2 Intelligence Platform: Application Server Administration Guide*.

1) Initiate the SAS Deployment Wizard.
2) Check **Configure SAS Software** and do not check **Install SAS Software** in the Select Deployment Type dialog box.
3) Select **Custom** in the Select Configuration Prompting Level dialog box.
4) Enter the configuration directory that is used for the original BI platform as the configuration directory in the Specify Configuration Information dialog box.
5) Select **SAS Stored Process Server**, **SAS Workspace Server,** and **SAS OLAP Server Configuration** in the Select Products to Configure dialog box.
6) Enter *SASApp_ZH* as the **SAS Application Server Context Name** in the SAS Application Server: Server Context dialog box.
7) Make sure that the ports that are used by the stored process server, workspace server, and OLAP server have not been occupied by other servers.

Modify the configuration file in order to enable the SASApp_ZH application server support Chinese.

8) Copy *sasv9.cfg* from <*SASROOT*>*\nls\ko to* <*SASROOT*>*\nls\zh* and name it as *sasv9.cfg*. If there is no *zh* folder under <*SASROOT*>*\nls*, then you need install SAS Foundation again and select Chinese as the support language. Update the configuration file. The following lines need be changed.

The Original Lines	The Updated Lines
`-SET SASCFG "C:\Program Files\SAS\SASFoundation\9.2\nls\ko"`	`-SET SASCFG "C:\Program Files\SAS\SASFoundation\9.2\nls\zh"`
`-DBCSLANG KOREAN`	`-DBCSLANG CHINESE`
`"!sasext0\ko\sashelp"`	`"!sasext0\zh\sashelp"`
`"!sasext0\ko\sasmsg"`	`"!sasext0\zh\sasmsg"`

9) Modify *sasv9.cfg* under <*SASCONFIG*>*/SASApp_ZH*. Note that *SASCONFIG* is the location of your configuration directory. For example, on Windows, 'C:\SAS\Config\Lev1' is an example for location of SASCONFIG.

Replace `-config "C:\Program Files\SAS\SASFoundation\9.2\sasv9.cfg"` with `-config "C:\Program Files\SAS\SASFoundation\9.2\nls\zh\sasv9.cfg"`.

Verify whether the SASApp_ZH Application server was configured successfully.

10) Restart service *SAS Object Spawner*.

11) Right-click the **Connection: SASApp_ZH-Workspace server** and Select **Test Connection** via SAS Management Console.

12) Right-click the **Connection: SASApp_ZH-Stored processed server** and Select **Test Connection** via SAS Management Console.

Enable SAS® Web Report Studio using SASApp_ZH-Workspace server to display a Chinese report.

13) Right-click **SASApp_KO – wrstemp** under **Data library Manager**-> **libraries** in the left pane of SAS® Management Console and select *properties*. In the **Assign** tab, select both *SASApp_KO* and *SASApp_ZH* as the selected servers.

14) Right-click **SASApp_KO** under **Server Manager** and select **properties**. In the **Authorization** tab, let Korean users have *ReadMetadata* permission and deny Chinese users *ReadMetadata* permission.

15) Right-click **SASApp_ZH** under **Server Manager** and select **properties.** In the **Authorization** tab, let Chinese users have *ReadMetadata* permission and deny Korean users *ReadMetadata* permission.

After the new SASApp_ZH application server is created, all stored process and cubes used to process Chinese data should be associated with the application server. Korean users will use SASApp_KO to display Web reports and Chinese users will use SASApp_ZH to display their reports.

Note[2]: In order to enable SASApp_ZH application server processing Chinese data normally, pay attention to the following items:

1) The paths to the Chinese reports, stored processes, and cubes should not contain Korean characters.
2) The names of SASApp_ZH application servers should not contain Korean characters.
3) The name of OLAP schema should not contain Korean characters.

DEPLOY A UNICODE SERVER TO SUPPORT MULTIPLE LANGUAGES

This paper only briefly discusses setting up a SAS Unicode Server. For complete details, see “Processing Multilingual Data with the SAS®9.2 Unicode Server” on the SAS Technical Support Web site.

Typically, a SAS BI Platform is deployed into three tiers: client tier, middle tier, and server tier. In order to support Unicode in a SAS BI platform, a Unicode server needs to be configured in SAS BI server tier to process multilingual data. Meanwhile, corresponding settings in the middle-tier and client tier need to be configured as well in a multilingual environment.

CONFIGURE A SAS UNICODE SERVER

A SAS Unicode Server is a SAS process running with UTF-8 encoding, which can handle multilingual data very well. The following steps show how to modify an existing SAS server to support Unicode in a Windows and UNIX operating system.

- **Windows**

  1) Stop all SAS services.
  2) Modify the *sasv9.cfg* file in the directory *<SASROOT>* (for example, *C:\Program Files\SAS\SASFoundation\9.2*).

     Replace the line*: -CONFIG "<SASROOT>\nls\en\SASV9.CFG"*

     with *-CONFIG "<SASROOT>\nls\u8\SASV9.CFG"*
  3) Start all SAS services.

- **UNIX**

  1) Stop all SAS services.
  2) Re-create a symbolic link from *<SASROOT>/sas* to *bin/sas_u8.*

```
$ cd <SASROOT> (for example, /usr/local/SAS/SASFoundation/9.2)

$ rm sas
```

[2] This note also applies to the SASApp_KO application server that is processing Korean data.

```
$ ln -s bin/sas_u8 sas
```

3) Start all SAS services.

Users can also install a SAS server as Unicode server from scratch by the SAS Deployment Wizard using the following steps.

- **Windows**

1) Initiate the SAS Deployment Wizard to do an installation without configuration.
2) Modify the sasv9.cfg file in the directory *<SASROOT> (for example, C:\Program Files\SAS\SASFoundation\9.2).*

   Replace the line: *-CONFIG "<SASRoot>\nls\en\SASV9.CFG"*

   with *-CONFIG "<SASRoot>\nls\u8\SASV9.CFG"*

3) Initiate the SAS Deployment Wizard again to do the configuration.

- **UNIX**

1) Initiate the SAS Deployment Wizard to do an installation without configuration.
2) Re-create a symbolic link from *<SASROOT>/sas* to *bin/sas_u8.*

```
$ cd <SASROOT> (for example, /usr/local/SAS/SASFoundation/9.2)

$ rm sas

$ ln -s bin/sas_u8 sas
```

3) Initiate the SAS Deployment Wizard again to do the configuration.

ENABLE LANGUAGE SWITCHING ON THE SAS SERVER SIDE

SAS 9.2 provides a new feature called ***Language Switching***, which allows users to view SAS messages in another language using the Language Switching feature in the Unicode server. Users can access the Language Switching feature with the LOCALELANGCHG system option. If LOCALELANGCHG is enabled, then the value of the LOCALE system option determines the language for procedure output, user interface elements, and ODS fonts. If LOCALELANGCHG is disabled, then messages appear in the language that is set during startup. This feature is supported in the Unicode server.

For more information about language switching, see "Changing language during a SAS session" on the SAS Technical Support Web site.

SWITCH THE LANGUAGE IN THE CLIENT

You can use the localized versions of client applications to connect to the Unicode server. The Unicode server can handle it smoothly. Sometimes, you might want to change the locale of the client application. For example, if you have SAS Management Console installed on Windows with a Chinese locale, and for some reason, you want to launch SAS Management Console in English , then you need change the locale of SAS Management Console. The detail steps on how to do this are listed below.

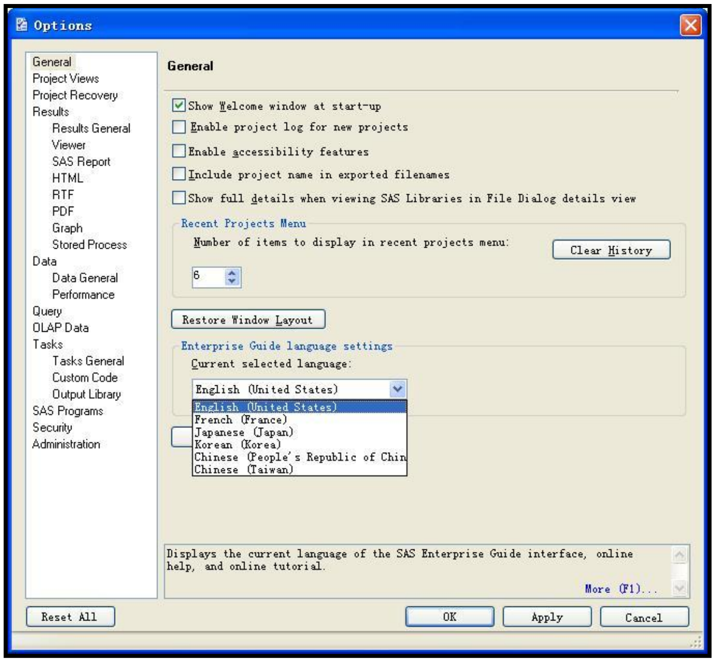
SAS client applications are divided into three types: Windows (SAS® Enterprise Guide®), Java (SAS Management Console, SAS® Data Integration Studio, and so on) and Web (SAS® Information Delivery Portal, and so on). There are different ways to change the language setting for each type of SAS client.

1. SAS Enterprise Guide

1) Start up the SAS Enterprise Guide 4.2.

2) Click **Tools** in the menu bar and click **Options**.

3) Select the language from the list.



General

☑ Show Welcome window at start-up
☐ Enable project log for new projects
☐ Enable accessibility features
☐ Include project name in exported filenames
☐ Show full details when viewing SAS Libraries in File Dialog details view

Recent Projects Menu
Number of items to display in recent projects menu: 6

Enterprise Guide language settings
Current selected language: English (United States)

English (United States)
French (France)
Japanese (Japan)
Korean (Korea)
Chinese (People's Republic of Chin
Chinese (Taiwan)

Displays the current language of the SAS Enterprise Guide interface, online help, and online tutorial.

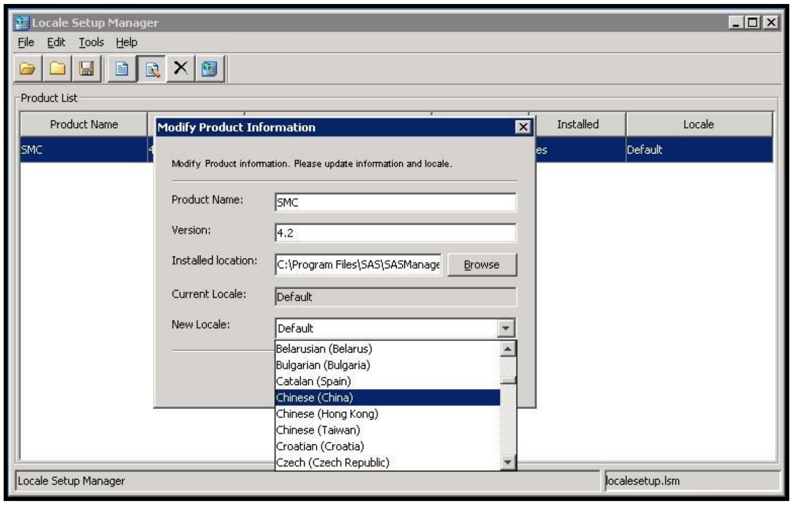
Figure 6. Change Language Settings in SAS Enterprise Guide

4) Restart SAS Enterprise Guide 4.2.

2. Java Clients

All of SAS Java clients’ default language settings can be changed using the Locale Setup Manager. This tool only supports Windows platforms.

In this case, you will set the locale for the SAS® Management Console to simplified Chinese.



Product List

Product Name: SMC

Installed: es

Locale: Default

Modify Product Information

Modify Product information. Please update information and locale.

Product Name: SMC

Version: 4.2

Installed location: C:\Program Files\SAS\SASManage

Current Locale: Default

New Locale: Default

Belarusian (Belarus)
Bulgarian (Bulgaria)
Catalan (Spain)
Chinese (China)
Chinese (Hong Kong)
Chinese (Taiwan)
Croatian (Croatia)
Czech (Czech Republic)

Locale Setup Manager

localesetup.lsm

Figure 7. Change Language Settings in the Locale Setup Manager

3. Web Browser-Based Clients

The SAS® Web Report Studio, SAS® Web OLAP Viewer, SAS® Information Delivery Portal, and SAS® BI Dashboard language settings are controlled by the Web browser.

For example, to change the language setting in Internet Explorer, follow this procedure:

1) Open Internet Explorer.

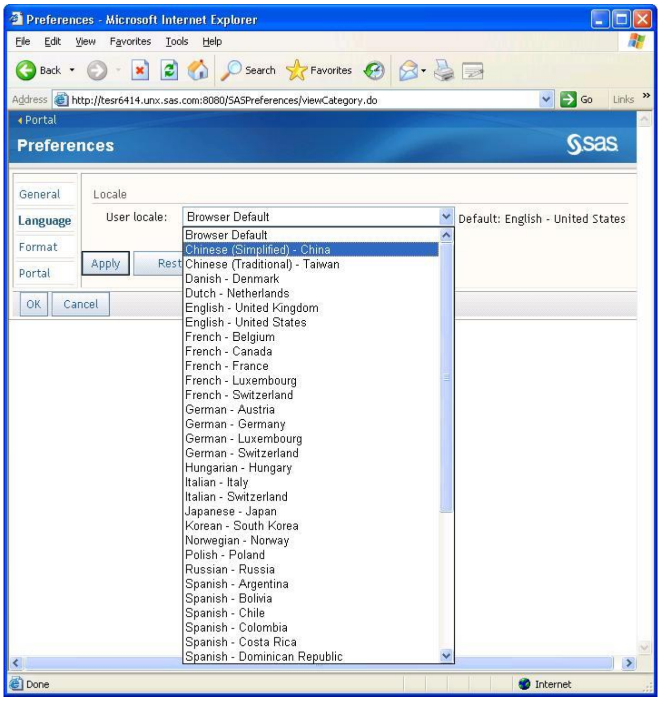
2) Select **Internet Options** from the **Tools** menu.

3) On the **General** tab click **Languages.**

4) Select the required language.

5) Click **OK.**

For the SAS Information Delivery Portal and SAS BI Dashboard, you can also change the language by setting **Options->Preference->language** in the **Portal** menu.



Preferences - Microsoft Internet Explorer

Address http://tesr6414.unx.sas.com:8080/SASPreferences/viewCategory.do

Preferences

Locale

User locale: Browser Default

Default: English - United States

Figure 8. Language Setting in the SAS Information Delivery Portal

HELPFUL HINTS AND TIPS

This article has introduced how to install and configure the SAS BI platform in a specified language, how to add a SAS application server to support an additional language, and how to configure SAS BI platform with a Unicode server. When you configure or use a SAS BI platform in a multilanguage environment, you often run into complex scenarios that are not covered in the above scenarios. Here are some hints and tips that might useful to solve many problems that you might encounter during configuring and using SAS the BI platform in a multilanguage environment.

CHANGE DATA SET ENCODING

We have an existing BI Platform using one encoding. Sometimes we want to change the BI platform to run on a new encoding. The existing data sets in the original encoding need be transcoded into the new encoding. Transcoding is the process of converting character data from one encoding to another encoding.

Transcoding is necessary when the encoding of data in the original location is different from the encoding of the data's destination. Target encoding should be a superset of source encoding. Otherwise, you will have some unreadable characters. Meanwhile, you should be aware of a data truncation issue that can occur during data transcoding. How to avoid data truncation will be detailed in the next section.

Example: Using the DATA step to change SAS data set encoding

```
LIBNAME source cvp ‘source path’;
LIBNAME target ‘target path’;

data target.targettable (encoding="utf-8");

set source.sourcetable;

run;
```

AVOID DATA TRUNCATION

The number of bytes for a character in one encoding can be different from that in another encoding. If the number of bytes for a character in a source encoding is smaller than that in a target encoding, then data truncation will occur. For example, when transcoding from a Central European Character set (wlatin2) to a Unicode encoding (such as UTF-8), the original variable lengths (in bytes) of the Central European Character set (wlatin2) characters might not be sufficient to hold the values. The result will be character data truncation.

When you specify the ENCODING= data set option, the encoding for the output data set might require more space than the original data set. In this situation, you can request that the character variable padding (CVP) engine expand character variable lengths so that character data truncation does not occur. The CVP engine is used in the above example.

Another method would be to calculate the size of the input buffer manually according to the bytes in the SAS session encoding and specify options with a large enough length setting when reading data. When you read data from an external fix-length record file, you should allocate the space for SAS variable length, which is called inner buffer size.

For example, there is a file encoded in EUC-CN where two Chinese characters reside. The two Chinese characters take four bytes in the file. If you read the file in a SAS Unicode environment with UTF-8 encoding, the column length of the result table should be set at least 6, otherwise data truncation will occur. Typically, one Chinese character takes 3 bytes in UTF-8 encoding.

Example: Specifying enough space when reading data from an external file

```
data target(encoding="utf-8");

Infile ‘c:\ dbcstext.txt’ encoding=’euc-cn’;

Input COL $ 6; /* the column length of result table should be set at least 6 */

run;
```

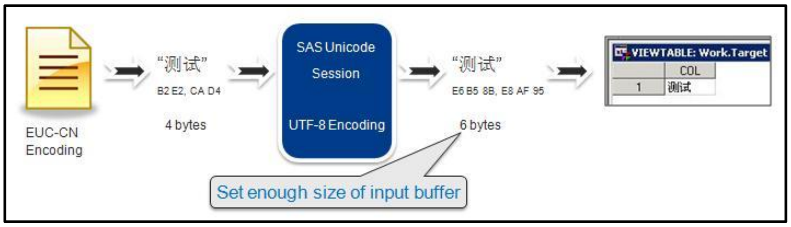


Figure 9. Specifying Enough Space when Reading Data from an External File

SETTING SYSTEM FONTS WITH X RESOURCE FILES ON UNIX

If the language that is used is Simplified Chinese, Traditional Chinese, Japanese, or Korean, then the SAS 9.2 Foundation might not have the correct font settings for the languages by default on UNIX. To ensure that the correct fonts are defined for the SAS System, you must add them to the X Resource files.

For example, Japanese X Resource template files that contain DBCS font settings are located in *<SASROOT>/X11/resource_files*, as follows:

./Resource_CDE.ja - for the CDE environment

./Resource_LNX.ja - for Linux

./Resource_Sun.ja - for Solaris

./Resource_DEC.ja - for Tru64 UNIX

./Resource_HP.ja - for HP-UX

./Resource_IBM.ja - for AIX

./Resource_RefIX.ja - for ReflectionX users

To apply the X Resources in these template files, copy the appropriate template to one of the following locations, renaming it to SAS (in all uppercase):

/usr/lib/X11/app-defaults (on most UNIX systems)

/usr/openwin/lib/X11/app-defaults (on Solaris)

For example, on an HP-UX system, you would use the following copy command:

```
$ cp <SASROOT>/X11/resource_files/Resource_HP.ja  /usr/lib/X11/app-defaults/SAS
```

CONCLUSION

The SAS BI platform can be deployed successfully in a specified language environment. In SAS 9.2, for some languages that the SAS Foundation doesn’t support, you need to update the configuration file manually to meet the requirements during deployment on a UNIX environment.

You can add a new SAS application server to support additional languages for existing BI platforms in SAS 9.2. If you want to maintain a central repository that can support a global set of languages, then you can deploy a BI platform with a Unicode server.

TERMINOLOGY

A **character set** is a collection of letters, symbols and punctuation marks used in a single language or in a group of languages.

A **LOCALE** reflects the language, local conventions such as data formatting, and culture for a geographical region. Local conventions might include specific formatting rules for dates, times, and numbers and a currency symbol for the country or region. Collating sequence, paper size, postal addresses, and telephone numbers can also be included in locale.

An **ENCODING** is the association of a unique numeric value with each symbol and punctuation mark in a character set. There are two groups or types of encodings: single byte character set (SBCS) encodings and DBCS encodings. SBCS encodings represent each character in a single byte. DBCS encodings require a varying number of bytes to represent each character. A more appropriate term for “DBCS” is multi-byte character set (MBCS). MBCS is sometimes used as a synonym for DBCS. SBCS encodings are limited to 256 possible characters. DBCS encodings can represent more than 256 characters.

UNICODE is the universal character encoding that supports the interchange, processing, and display of characters and symbols found in the world’s writing systems. Other character encodings are limited to subsets of all languages. Often the subsets are limited to a region (for example: Windows Latin 1 [WLATIN1] represents the characters of the US and Western Europe on Windows). UTF8 is the Unicode encoding form where characters are represented in sequence of one to four bytes in length.

National Language Support (**NLS**) is a set of features that enable a software product to function properly in every global market for which the product is targeted. The SAS System contains NLS features to ensure that SAS applications can be written so that they conform to local language conventions.

REFERENCES

SAS Institute Inc. 2009. *Configuration Guide for SAS® 9.2 Foundation for UNIX® Environments*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2010. *SAS® Deployment Wizard User's Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. *SAS® 9.2 Intelligence Platform: Application Server Administration Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. *SAS® 9.2 Intelligence Platform: Installation and Configuration Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. *SAS® 9.2 National Language Support (NLS): Reference Guide*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. *What's New in SAS® 9.2*. Cary, NC: SAS Institute Inc.

SAS Institute Inc. 2009. SAS Institute technical paper. "Processing Multilingual Data with the SAS® 9.2 Unicode Server." http://support.sas.com/resources/papers/92unicodesrvr.pdf.

SAS Institute Inc. 2008. SAS Institute technical paper. “Changing Language During a SAS Session.” http://support.sas.com/resources/papers/LocaleSwitching.pdf.

ACKNOWLEDGMENTS

The author wishes to express special thanks to Kansun Xia. He suggested a lot to this paper from the beginning and exerted lots of efforts to review and revise this paper. Also, we wish to thank the following people (listed in alphabetical order) who have taken time to review and contribute to this paper: Betty Joyner, Caroline Quinn, Elizabeth Bales, Huong Le, Manfred Kiefer, Masahiro Takagi, Jin Luo, Shin Kayano, and Steve Beatrous.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author:

Name: Holmes Niu
Enterprise: SAS Research and Development (Beijing) Co., Ltd
Address: 19/F North Wing, Central Office Tower, Junefield Plaza, No. 10 Xuanwumenwai Dajie Xuanwu District Beijing China, 100052
Work Phone: +86 10 8319-3631
Fax: (8610)6310-9130
E-mail: Holmes.Niu@sas.com