

Keep your eye on the code page, practical exam

 devblogs.microsoft.com/oldnewthing/20060301-14

March 1, 2006



Raymond Chen

The instructions that came with my new monitor are printed in several languages. One of those languages is Polish, or at least it would be Polish if... well, you'll see.

| UWAGA: Szczegó³owe informacje dla u^zytownika znajduj¹ siê na do³¹czonej p³ycie CD.

This is garbage. What they meant to write was

| UWAGA: Szczegó^lowe informacje dla u^zytownika znajduj^a siê na do^laczonej p^lycie CD.

(It means, "Note: Detailed user information is included on the CD.")

What went wrong? Why did all but one of the accented characters come out corrupted?

The company responsible for the instructions failed to keep their eye on the code page. The text provided to them by their Polish translator was in code page 1250 (Central Europe). If you hold up that code page diagram next to code page 1252 (Latin I), you can see where the problem is. The company took that string in code page 1250 and printed it in code page 1252. For example, the "ł" character is at position B3 in code page 1252; the character at that position in code page 1250 is "3".

I just hope their Polish customers can figure out what the text is supposed to say.

Okay, that was their practical exam. Here's yours:

I am running on a Spanish install of Windows 2003 and calling

```
WideCharToMultiByte(CP_OEMCP, WC_NO_BEST_FIT_CHARS,  
                    pwzStr, -1, pszStr, nRet, NULL, pbUsedDefault);
```

When I pass this Unicode string

```
D:\Documents and Settings\ABC\Configuración local\
```

it returns the multibyte string below and doesn't set `pbUsedDefault` :

```
D:\Documents and Settings\ABC\Configuraciϕn local\
```

Why isn't `pbUsedDefault` set?

In fact, I'm going to add two bonus questions: (1) How should this customer change their code to get the path name converted correctly? (2) Why? The answer to the second question is the important one.

Raymond Chen

Follow

