Subject: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Sat, 29 Nov 2008 15:08:23 GMT
View Forum Message <> Reply to Message

I try use GridCtrl with MySQL. On number column sorting work fine, but on String columns sorting very strange. How to setup right collating sequence in GridCtrl. Please Help

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by unodgs on Wed, 10 Dec 2008 07:23:57 GMT
View Forum Message <> Reply to Message

sergeynikitin wrote on Sat, 29 November 2008 10:08I try use GridCtrl with MySQL. On number column sorting work fine, but on String columns sorting very strange. How to setup right collating sequence in GridCtrl. Please Help
What you mean exactly. GridCtrl just uses Value comparator. So if value is String - String comparator will be used, if value is WString - WString comparator will be used and so on.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Wed, 10 Dec 2008 20:48:48 GMT
View Forum Message <> Reply to Message

Yes, I understand. I have found in sources ...

But if you are prompted for example, I would be grateful.

Thanks in advance.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Mon, 04 May 2009 10:22:00 GMT
View Forum Message <> Reply to Message

I propose to do in .t file a line like "ABCDEFG ...... abcdefg ..... vwxyz ...." with the correct sequence of characters, then when you switch the language will work correctly sort and compare strings.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by mr_ped on Tue, 05 May 2009 07:07:08 GMT
View Forum Message <> Reply to Message

sergeynikitin wrote on Mon, 04 May 2009 12:22I propose to do in .t file a line like "ABCDEFG ...... abcdefg ..... vwxyz ...." with the correct sequence of characters, then when you switch the language will work correctly sort and compare strings.

Will not work for all languages in this simple form. For example cz and sk have special character "ch" (between "h" and "i" in alphabet). The funny thing is, that this correct local sorting always annoys hell out of me and I absolutely hate it (and never ever needed it). Then again, most of the time I use English anyway.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Tue, 05 May 2009 12:23:04 GMT
View Forum Message <> Reply to Message

For example this might look like this:

deDE_Collating_Sequence("ABCDEFGHIJ......abcdefghij.....")

And then switched the language, both switches and sorting mechanism.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by mr_ped on Wed, 06 May 2009 07:01:14 GMT
View Forum Message <> Reply to Message

but "ch" are two chars, "c" and "h", there is never "ch" character in fonts, because it was always written from the two. Maybe I just don't understand you, and there's some hidden way how to

the simple lines you posted.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Wed, 06 May 2009 07:28:56 GMT
View Forum Message <> Reply to Message

No, Sorry. I'm not quite clear.

This is a problem with the editing of Unicode characters in the forum.

J - it was an attempt to insert the letter, with emphasis. (Precisely what it turned out)

That is, there is a letter designating the ch or just a couple of characters in the alphabet? I do not understand.

!!!
But in any case to establish the sort in the same line used MySQL Collating Sequence, which has consistently recorded the entire alphabet, letter for letter, usually starting with English letters and ending with the national.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by mr_ped on Wed, 06 May 2009 08:39:24 GMT
View Forum Message <> Reply to Message

sergeynikitin wrote on Wed, 06 May 2009 09:28That is, there is a letter designating the ch or just a couple of characters in the alphabet? I do not understand.

That's the tricky part about it. In Czech alphabet "ch" is single letter, and for example dictionaries are sorted by it, i.e. you have:
banik, cislo, frajer, hroch, chyba, ikra
(notice position of "cislo" and "chyba")

But it is written from regular "c" and "h" letters.

This did lead to a situation where in the digital land the "ch" is always represented by regular "c" and "h" characters, i.e. the single "ch" letter is formed in digital data by regular two characters. This means that in correct cs/sk sort you have to compare not just single character (bytes), but in case you hit "c", you have to look ahead for next character, if it is "h", you have special case.

Unless you use special substitution character for every "c" "h" pair before sort, and then replace it back, you can't cover this behavior by simple order-table.

edit:
If my personal opinion would count, I would discard "ch" immediately from our alphabets, so all the "ch" words would be created from regular "c,h" pairs, i.e. pretty much the current situation except weird sort and weird letter count.
Then again, if my personal opinion would count, there would be no Russian alphabet at all, because I said in school once at Russian lesson: "a nation which is unable to learn+use English 26 letter alphabet will vanish in future". (I was 10 years old at that point and I was so glad ASCII standard existed, so I hope you will understand it a bit)
Since then I admit the utf-8 may be remedy of current status, still if everyone would switch to ASCII 20 years ago and discard all funny chars (and we got much more of them then just "ch", plenty of accented ones too), I bet the current SW would be much smoother ride and lot more cheaper then it is....

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by sergeynikitin on Wed, 06 May 2009 09:26:01 GMT
View Forum Message <> Reply to Message

I do not understand what it contradicts.

If there is a complex situation with a number of scripts, rather hard-ASCII heritage. That does not mean that you do not need to enter the normal Unicode tools.

I understand that the above problem is not everywhere, but only in certain applications.

In any case, the MySQL and PostfreSQL solve this problem the way that I proposed for U++.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by Mindtraveller on Sun, 10 May 2009 06:26:15 GMT
View Forum Message <> Reply to Message

mr_ped wrote on Wed, 06 May 2009 12:39Then again, if my personal opinion would count, there would be no Russian alphabet at all, because I said in school once at Russian lesson: "a nation which is unable to learn+use English 26 letter alphabet will vanish in future". (I was 10 years old at that point and I was so glad ASCII standard existed, so I hope you will understand it a bit)Excuse me for little offtopic.
National features topic is a delicate one. It all depends on a balance. You may always choose between simplicity and possibilities. For example, when we talk about language, we all know about international, or basic English. It is a subset of common English language (rather tiny comparing to Britain English). Basic English is an example of simplicity. But it has no way to say anything beyond the scope of anything common (i.e. you can`t translate any national object or any specific thing, feeling except of those predefined subset).
Contrary each national language has unique features but less simple. This unique information is commonly contained in writing system itself and turning it into basic English alphabet will petrify 90% of this information.
Let`s take Chinese as an example. It has a number of delicate definitions which have no equivalents not only in basic English but in many other languages too. More of that. Hieroglyphic system contains information on national mentality. For example there`s hieroglyph "woman". And hieroglyph "woman" under "roof" means calm, peace of mind. Hieroglyph consisting of two "woman" hieroglyphs mean "scandal" (IMO it`s cool!). Even more, hieroglyph consisting of three "woman" means even greater scandal or debauch (cool!). I consider this writing system very logical and intuitive. And it is certainly worth keeping.
Some languages have 5-10 words describing love (with delicate differences each other), or even many-many words describing beauty of nature.
When we talk about cyrillic writing system we have even more delicate information layer based on radicals and other parts of a word. More interesting is these radicals are common between most of slavonic languages. For example when you know 2 or 3 of them, you may possibly translate text from i.e. Polish or Chech. It is closely related to cyrillic letters which introduce logical system of radicals (with their geographic/historic variations) keeping some kind of unity between any slavonic language (and mentality as consequence).
As a conclusion I`d say that basic English is a good tool for me if I want anyone from outside to understand me. But each nation has something unique which is closely related with their language and writing system. IMO it is worth keeping especially comparing to the alternative of petrifying process of thinking with switching to basic English. Again, Britains themselves use English which is more complex and rich than basic one.

Subject: Re: How can I setup right collating sequence in GridCtrl?
Posted by mr_ped on Sun, 10 May 2009 09:40:03 GMT
View Forum Message <> Reply to Message

I hope you didn't miss the "10 year old" part of my comment.
I did evolve since then a bit.
Still interesting reading, thank you.