Subject: Speller routine
Posted by BetoValle on Mon, 11 Jan 2021 14:10:09 GMT
View Forum Message <> Reply to Message

when the user activates spellcheck in the RichEdit editor, depending on the configured language, the system searches for the .udc file.
The point is that here in my country, from time to time, the dictionary grows.
In this way it is necessary (in my case) that the user can have access (in a friendly way) to the words in this dictionary.
(in my delphi applications I used it in the component editor trichedit / www.trichview.com / to update the custom dictionary that was loaded from an editable database).
(a) Is there a structure defined in U ++ / RichEdit allowing the dictionary to be editable?
(b) What is the internal structure of a .udc file (probably compressed)? (I think this would be the easiest way to adapt).

Thanks!

Subject: Re: Speller routine
Posted by BetoValle on Tue, 12 Jan 2021 19:27:50 GMT
View Forum Message <> Reply to Message

Could someone give me a light? thanks

Subject: Re: Speller routine
Posted by BetoValle on Wed, 13 Jan 2021 13:14:30 GMT
View Forum Message <> Reply to Message

There is a situation to expand the dictionary, which is to create a file similar to that of spell (.udc), but with the extension ".usp" saved inside the application folder, where new words are added (edit a text file where each new word is on a line). After this addition, the user will have to close the application to reestablish this update.
Couldn't a routine be created in this situation to read the '.usp' file again as a "SpellerRefresh"? So it would not be necessary to close and load the program again (I don't know how to do this, because the structure is complex ... Maybe someone from there with good will can do it).
Well as I understand it, there is a lack of a routine to remove from the '.udc' file, the words that are eventually discarded by an orthographic reform (this exists!)

Subject: Re: Speller routine
Posted by mirek on Fri, 22 Jan 2021 12:26:23 GMT
View Forum Message <> Reply to Message

BetoValle wrote on Mon, 11 January 2021 15:10when the user activates spellcheck in the RichEdit editor, depending on the configured language, the system searches for the .udc file.

The point is that here in my country, from time to time, the dictionary grows.
In this way it is necessary (in my case) that the user can have access (in a friendly way) to the words in this dictionary.
(in my delphi applications I used it in the component editor trichedit / www.trichview.com / to update the custom dictionary that was loaded from an editable database).
(a) Is there a structure defined in U ++ / RichEdit allowing the dictionary to be editable?
(b) What is the internal structure of a .udc file (probably compressed)? (I think this would be the easiest way to adapt).

Thanks!

a) no
b) uppbox/MakeSpellScd Despite the name, this produces .udc format...

Subject: Re: Speller routine
Posted by BetoValle on Fri, 22 Jan 2021 13:23:07 GMT
View Forum Message <> Reply to Message

Hi
I created a file pt_br.usp containing words that the file pt_br.udc did not contemplate, I put it in the root of the application and it works. So in topic (a) by the written sources it allows this functionality.
In the file pt_br.usp, it makes no sense for me to write
a variation of the same word ... example:
anything
anyTHING
Anything ... etc ...
To solve this, inside the function "spellWordRaw" (Speller.cpp) every line containing "return f-> user.Find (wrd)> = 0" could be replaced by the return of a function below "return SpellWordRawMeu(f,wrd);" :

```
bool SpellWordRawMeu( Speller *f, const WString& wrd ){
 String t1 = ToUtf8(wrd);
 String t2 = ToUtf8(ToLower(wrd));

 WString w1 =t1.ToWString();
 WString w2 =t2.ToWString();

 boolean b1 =  f->user.Find(w1) >= 0;
 boolean b2 =  f->user.Find(w2) >= 0;
 if(b1 || b2){
  return true;
 }else{
  return false;
 }
}
```

in preliminary tests, it worked!
So, in order not to disturb the others written in the code that already works, externally, the user of my country could keep the file with udc extension without content, and create a partial file with usp extension, with all lowercase words (the which is not difficult to get here / I already did this to test). So a routine can easily be written into a routine for the user to update the usp file if necessary.

I searched on google about MakeSpellScd, but the only 2 references fall into the discussion where you already know.

Thanks
