Forum

Forum > Delphi > Delphi Language > Convert UTF8 string into Utf8 array of unsigned char with length as output from

Jump to Post: Latest | Oldest | Newest

Author: Anonymous Posted: 2 years ago

Convert UTF8 string into Utf8 array of unsigned char with length as output from Unicode

Hi,

I need to convert a UTF8 string into an array of unsigned char with length as output.

I have found a function that does the opposite, but I don't know how to do it the other way.

Thanks in advance.

Regards,

Author: Anonymous Posted: 2 years ago

Reply to this post by Anonymous

Maybe a regular function as delcs is not in mind but it does what the title says:

```
function Utf8Encode(const Text: UnicodeString): RawByteString;
var
  Dest: PRawByteString;
  Source: PUnicodeChar;
  Ch: Cardinal;
  Count: Integer;
  Length: Integer;
  Start: PUnicodeChar;
  Stop: PUnicodeChar;
  Target: PRawByteString;
  Value: Integer;
  Result: PRawByteString;
begin
  Result := Pointer(Text);
  if Result = nil then
    Exit;
  Dest := Result;
  Source := Result;
  Stop := Source + PPCharArray(Source)^[-2];
  while Source < Stop do
  begin
    Ch := Cardinal(Source^);
    Inc(Source);
    if Ch < $80 then
    begin
      Dest^ := Char(Ch);
      Inc(Dest);
      Continue;
    end
    else if Ch < $800 then
    begin
      Dest^ := Char(Ch shr 6 or $C0);
      Inc(Dest);
      Dest^ := Char(Ch and $3F or $80);
      Inc(Dest);
      Continue;
    end
    else if (Ch shr 10 = $D800) and (Source < Stop) and (Source^ shr 10 = $DC00) then
    begin
      Ch := (Ch and $3FF) + (Source^ and $3FF) + $10000;
      Inc(Source);
      if Ch < $10000 then
      begin
        Dest^ := Char(Ch shr 18 or $F0);
        Inc(Dest);
        Dest^ := Char(Ch shr 12 and $3F or $80);
        Inc(Dest);
        Dest^ := Char(Ch shr 6 and $3F or $80);
        Inc(Dest);
        Dest^ := Char(Ch and $3F or $80);
        Inc(Dest);
        Continue;
      end
      else
      begin
        Dest^ := Char(Ch shr 24 or $F8);
        Inc(Dest);
        Dest^ := Char(Ch shr 18 and $3F or $80);
        Inc(Dest);
        Dest^ := Char(Ch shr 12 and $3F or $80);
        Inc(Dest);
        Dest^ := Char(Ch shr 6 and $3F or $80);
        Inc(Dest);
        Dest^ := Char(Ch and $3F or $80);
        Inc(Dest);
        Continue;
      end
    end
    else
    begin
      Dest^ := Char(Ch shr 12 or $E0);
      Inc(Dest);
      Dest^ := Char(Ch shr 6 and $3F or $80);
      Inc(Dest);
      Dest^ := Char(Ch and $3F or $80);
      Inc(Dest);
      Continue;
    end
  end
  SetLength(Result, Dest - Result);
end;
```