

同じに見えるのに一致しない？ 見えない文字を調べよう

ワカサギさんたちと、文字の中身と可視化の限界を確かめる

Copyright © 2026 LWP 山中 一弘

本資料は、出典を明記いただければ、商用・非商用を問わず、ご自由に複製・改変・再配布していただけます。なお、著作権表示は改変せず、そのまま記載してご利用くださいますようお願いいたします。

ストーリー

文字比較の記事をまとめているワカサギさんが、シャケモンへ相談します。画面では同じように見えるのに、比較すると違うと出る文字列があるのです。中の文字を表示してみると、原因だけでなく、調べ方にも条件があることが分かってきました。途中からアユさんも加わります。

会話と場面は架空の説明例です。対象はWindows版Excel VBAとTakiLib 3.5.4。図の文字列は仕組みを示す模式図であり、実際の画面表示はソフトやタブ幅によって変わります。

1 同じに見えるのに、一致しない

空いている部分が普通のスペースなのか、タブなのか。見た目だけでは判別しにくいことがあります。違いを直そうとする前に、入力に何が含まれているかを確認めます。



図1 | 空いて見える場所にも、文字が含まれている。

2 タブの居場所が見えた

`A`、タブ、`B`の順に並ぶ文字列を、`textHiddenCharToHex`で観察します。囲みを`[]`と指定すると、タブが`[0x09]`に置き換わり、位置を読み取れます。`0x`は、ここでは16進数を示すための表示です。



図2 | 今回の基本例は、タブ1文字の可視化。

3 観察用の結果は、別に置く

`[0x09]`は、元のタブそのものではありません。人が読むための複数の文字です。観察用の表示を元データへ戻すと、中身が変わります。元データは残し、確認用の結果を別の場所に置きます。



図3 | 可視化は観察の手段。修復や復元を保証する処理ではない。

4 改行にも種類がある

CRとLFは、それぞれ異なる制御文字です。CRLFは2文字の並びです。TakiLibの`textCanonicalNewline`は、CRLFと単独CRをLFへそろえます。どの並びをどの形式へ変えるかを明確にして使います。



図4 | この処理でそろえる先はLF。

5 置換する順番が大事

CRLFのCRだけを先にLFへ変えると、もともとのLFも残り、LFLFになります。現行実装はCRLFの組を先に1つのLFへ置換し、その後に残った単独CRをLFへ変えています。

5

置換する順番が大事

✕

CRだけ先に置換

CR

LF

➡

LF

LF

改行が
2つに

✓

CRLFを先に置換

CR

LF

➡

LF

残った単独CRも
LFへ



順番だけで、
結果が変わる
んですね。



2文字の組を、
先に置き換えます。

図5 | 複合した並びを先に扱うことで、改行の二重化を防ぐ。

6 見えない文字を、全部探せる？

この可視化関数は、不可視文字の一覧を調べているわけではありません。現行実装では`AscW`の値が`&H20` (10進数の32) 以上なら、その文字をそのまま返します。普通のスペース、NBSP (改行しないスペース)、ゼロ幅スペースは、この記事の処理では可視化されず残る例です。



図6 | 「この例で見つかった」を「全部見つかる」へ広げない。

7 見える全角文字まで、変わった

逆に、見える文字が置換側へ入る場合もあります。Windows VBAで全角の`A` (U+FF21)を`AscW`へ渡すと、符号付き16ビット値の`-223`になります。負の値は32以上という条件を満たさないため、この関数では置換側へ進みます。文字の見た目を判定した結果ではありません。



図7 | 図のAは半角Aではなく、全角A。

8 末尾2桁では、元に戻せない

置換側では、16進表記の末尾2桁を使います。全角`A`の例でも末尾の`21`だけが表示に使われ、既定の囲みなしでは`0x21`になります。この短い表示だけを見て、元の文字や文字コードを特定することはできません。



図8 | 0x21という表示を、元の文字コードそのものだと解釈しない。

9 観察と修正を、分けよう

改善を考えるなら、符号付きの値を符号なしとして扱う処理と、どの文字を可視化の対象にするかという分類を分けます。たとえば`AscW(xch) And &HFFFF&`という符号なし化だけで、すべての不可視文字を検出できるようになるわけではありません。



図9 | 今回は改善の考え方まで。ライブラリ本体は変更しない。

まとめ

見た目が同じでも、文字の並びは違うことがあります。可視化はその観察に役立ちますが、表示結果を元データへ上書きせず、検出範囲と情報の欠落も確かめます。改行をそろえるときも、置換する順番まで含めて処理を理解することが大切です。

今回の仕掛け

TakiLib 3.5.4の`textHiddenCharToHex`は、各文字の`AscW`値を調べて表示を作ります。
`a\_quote:=""`なら、可視化した箇所だけを角括弧で囲みます。`a\_alt\_char`を省略した例では16進表記を使います。

`textCanonicalNewline`は、まずCRLFをLFへ置換し、次に残ったCRをLFへ置換します。可視化関数と組み合わせると、実際の改行を画面に出す代わりに、その並びを記号で確認できます。

今回のソースコードのサンプル

サンプルコード タブ・改行・全角文字を比べる

```
01 Option Explicit
02
03 Public Sub Test_Main()
04     Dim s As String
05     s = "A" & vbTab & "B"
06     Debug.Print taki3lib.textHiddenCharToHex(s, a_quote:=""")
07
08     s = taki3lib.textCanonicalNewline("A" & vbCrLf & "B")
09     Debug.Print taki3lib.textHiddenCharToHex(s, a_quote:=""")
10
11     s = ChrW(&HFF21&)
12     Debug.Print AscW(s)
13     Debug.Print taki3lib.textHiddenCharToHex(s)
14 End Sub
```

コードの説明

最初にタブを含む文字列を可視化します。次にCRLFを正規化し、その結果を可視化します。最後に、入力時の半角・全角の混同を避けるため`ChrW`で全角Aを作り、符号付きの値と現行関数の表示を確認します。元のセルを書き換える処理はありません。

入力・概要

Windows版Excelの確認用ブックへ、公開安定版TakiLib 3.5.4の通常CP932版`taki3lib.bas`をVBEからインポートした環境を前提にします。AI読解用の`taki3libUTF8.bas`はインポートしません。別の標準モジュールへのマクロを置き、イミディエイトウィンドウで出力を確認します。

結果

ソースと仕様に基づく期待値は、順に`A[0x09]B`、`A[0x0A]B`、`-223`、`0x21`です。掲載マクロは今回Excelで実行検証していません。会話中の例も、今回の実測結果としては扱っていません。

サンプルコード 可視化関数の判定箇所を読む

```
01 xcode = AscW(xch)
02 Select Case True
03     Case xcode >= &H20
04         xtmp = xtmp & xch
05     Case XC_IS_ALT_MISSING
06         xtmp = xtmp & textFormat("{}0x{}", _
```

```

07      xquote1, Right("0" & Hex(xcode), 2), xquote2)
08  Case Else
09      xtmp = xtmp & textFormat("{}{}", _
10      xquote1, a_alt_char, xquote2)
11 End Select

```

コードの説明

直接呼び出している`textHiddenCharToHex`から、文字ごとの判定部分を抜粋しました。`xch`は処理中の1文字、`XC\_IS\_ALT\_MISSING`は代替表示の引数を省略したかどうかです。32以上なら元の文字を保持し、それ以外で代替表示を省略した場合には`Right(..., 2)`で末尾2桁を使います。掲載幅に合わせて行継続を加えています。この抜粋だけでは実行できず、ブックへ追加するコードではありません。

サンプルコード 改行正規化関数の原典抜粋

```

01 Public Function textCanonicalNewline(a_text)
02     If TK3_ASSERT Then _
03         Call langAssert(hasCStrAffinity(a_text), TK3_ERROR_ARGUMENT)
04     Dim x_text:
05         x_text = Replace(a_text, vbCrLf, vbLf)
06     textCanonicalNewline = Replace(x_text, vbCr, vbLf)
07 End Function

```

コードの説明

直接呼び出す関数の本文から、説明コメントと空行を除き、掲載幅に合わせて行継続を加えた読解用抜粋です。CRLFを先に置換する順序が確認できます。導入済みTakiLibと重複するため、追加インポートはしません。

出典メモ

- 元ネタ: TK-050「見えない文字を見えるようにする」(2026年9月9日構成再精査)。
- TakiLib公開安定版3.5.4: `textHiddenCharToHex`、`textCanonicalNewline`。2026年9月27日に公開版と該当ソースを確認。
- Microsoft Learn「[Asc function](#)」。AscWのUnicode文字コードに関する説明を参照。2026年9月27日確認。