Correction of SEX value based on HUSB and WIFE references

Functionality

In GEDCOM files you will find sometimes incorrect information about the gender of individuals.

The program checks for every family (FAM) record the HUSB and WIFE references and its values stored for SEX and allows a correction of this SEX value resp. an exchange of the HUSB and WIFE references. Therefor only those persons who are referenced as HUSB or WIFE can be checked. The corrected data can be stored in a new file with the filename of the original file and an appended "_shw" (xxx.ged will become xxx_shw.ged). The original file remains unchanged. The modified new file is saved in the folder of the original file. In addition, a log file is created that can be opened with any text editor. The name of the log-file is composed by the name of the ged-file with an appended "_shw.log" (xxx.ged will become xxx.ged_shw.log). The log-file contains:

- Statistics about the analysis.
- The corrections with indication of the type, the record numbers and the names of the persons concerned.

Description

After starting the program, above main screen will show up.

After loading a ged-file (by the [Symbol with an open folder]) the analysis can start after pushing [Start]. This will be done in 2 steps:

- Reading the whole file into storage, preparing the data and analysis of discrepancies with respect to the assignment of the gender by the SEX tag. This will be done by the program automatically. After analysis following will be listed at "Findings":
  - The number of differences found and the differences still open.
  - The 1st found difference not yet processed will be listed at the text field as 2 or 3 line information with the details:
    - FAM @Fxx@ Id
    - HUSB @Ixx@ Id, his name and sex value

- WIFE @Ixx@ Id, her name and sex value

The 1st line is always the FAM record number concerned. The 2nd and 3rd line, if needed, contains the information for HUSB and/or WIFE, as wll as the recommended corrections. 2 lines will only be displayed, if HUSB or WIFE is not included in the family record. In above example:

```
FAM @F6041@
1 HUSB @I36230@ = SEX "U" <- "M" [ Nn /Höllermann/]
```

- The correction of the differences found. The program offers for each discrepancy the possible changes and the user must decided which one should be performed. These are 4 buttons available. These are:
  - [HUSB <> WIFE] - If 2 lines are displayed the text "HUSB" will be converted to "WIFE" and vice versa, in case of 3 lines the reference IDs of HUSB and WIFE will be exchanged.
  - [Change SEX of HUSB] - This will convert for the person referenced as HUSB the actual character to the proposed character (in our example the "U" will be replaced by an "M").
  - [Change SEX of WIFE] - This will convert for the person referenced as WIFE the actual character to the proposed character.
  - [Next] - This will store the actual values and displays the next finding, independent if corrections have been done. A switch back is not possible. Skipped or not completely processed discrepancies can be processed in another follow-on run.

  After the differences of a finding are finally processed, a green "Ok" will show up above the text field. In case, the option "**Ok = Auto Next**" is selected, the program switches automatically to the next finding. without pushing the [Next] button.
  After correcting a discrepancy, the color of the [Save] button is displayed as a warning in "orange". This means that data have been changed but not yet stored in the modified file. **Note**: It is not necessary to save after each modification.

- Upon completion of processing, the modified data must be saved by [Save] to get the data into the modified file. This is especially required in case of an "orange" button. When loading a new ged-file and during closing the program, a check will be done by the program and, if necessary, a corresponding confirmation prompt will show up.

Additional buttons

- [Symbol with an open folder] Load the file resp. the folder.
- [Orange-blue symbol] Starts, if installed, Utility program "WinMerge" with both ged-files and performs a file comparison.
- [Edit] Pushing by the left mouse button opens the ged-file by the text editor for viewing. Pushing the right mouse button starts "GedShow" with the ged-file.
- [∧] Moves the modified file to the ged-file to be analyzed again.
- [Save] Saves the modified file. The color is "orange" after a modification, as long as these are not saved.
- [New] Clears all definitions.
- [Info] Shows the folder where the ini-file is stored (contains all settings at the last "Close").
- [Start] Starts the analysis.
- [Edit Log] Opens the log-file by the text editor.
- [Close] Writes the ini-file with the settings, opens, if necessary a "Save" dialog and closes the program. Closing the program by the little "x" at the upper right corner of the screen will not write the ini-file and will not start a "Save" dialog.)

Examples and explanations for discrepancies and their correction

As possible actions are available for below example:

- "HUSB <> WIFE" - this would exchange for FAM @F60013@ the reference numbers of HUSB and WIFE.
- "Change SEX of HUSB" - this would convert for INDI @I148088@ the text "F" for SEX into "M"
- "Change SEX of WIFE" - this would convert for INDI @I142351@ the text "M" for SEX into "F"

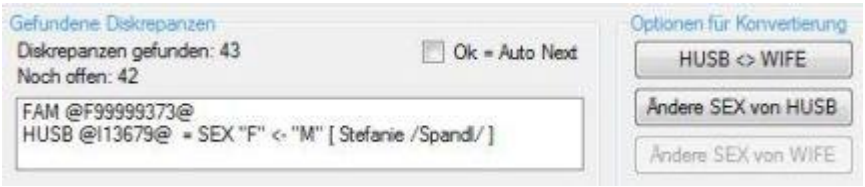
The first action is certainly to be recommended here, because an inadvertent mix-up happened when entering HUSB and WIFE data. This is particularly noticeable by watching the given name.

As possible actions are available for below example:

- "HUSB <> WIFE" - this would exchange for FAM @F6452@ the reference numbers of HUSB and WIFE.
- "Change SEX of WIFE" - this would convert for INDI @I48426@ the text "U" for SEX into "F"

The latter would certainly be recommend here - watch the given name.

As possible actions are available for below example:



- "HUSB <> WIFE" - this would convert for FAM @F99999373@ the text HUSB into WIFE
- "Change SEX of HUSB" - this would convert for INDI @I13679@ the text "F" for SEX into "M"

The first action is certainly to be recommended here, because the given name points to a correct WIFE entry.

As possible actions are available for below example:

- "HUSB <> WIFE" - this would convert for FAM @F6041@ the text HUSB into WIFE
- "Change SEX of HUSB" - this would convert for INDI @I36230@ the text "U" for SEX into "M"

Here, the decision is more difficult, since it cannot be detected accurately by the non-unique given name, whether it is here a HUSB or WIFE. The latter would be well recommended. A wrong decision would probably be detected and corrected by another run.

Specifics of the processing

To find discrepancies with respect to the stored SEX character, the program analyzes the FAM records from top to bottom. Single individuals may show up several times by e.g. multiple marriages. The discrepancies found will be stored, in case of multiple found individuals always the last found discrepancy. Because a person can occur several times, both with a correct and with a false SEX value (once as HUSB, once as WIFE), special caution is necessary in deciding the appropriate action. If there is a proposal for HUSB to be a "F", although there appears a male given name, this should probably not be changed. At this point the "M" is the correct value, but not at an other discrepancy for this person - or vice versa.

After finishing all modifications the 1st time (don't forget to [Save] the data !!) the modified file "..._ shw.ged" should be loaded again for a 2nd and possibly 3rd run (can be done by the [∧] button), until no more discrepancies are found. Then a final examination should be performed by the "GSP Validator" and errors still found should be manually corrected by using the genealogy program or a text editor.

WARNING: The modified file will be saved only by pressing the **[Save]** button.

Extract of a Log-file

```
Created at 04.05.2015 15:40:45 - CorSexHusbWife Version: 0.9.1
Base: "C:\Ofb\1\ftm\nikolsburg_2015-03-21_u.ged" from 29.04.2015 20:11:24

04.05.2015 15:40:58: Reading data completed
480.252 Total lines read

04.05.2015 15:40:58: Analysis completed
Total Differences found: 42

Modifications done:
 -> "HUSB <> WIFE": HUSB @I4776@ -> WIFE @I4776@  [ Maria /Swoboda/ ]
 -> "HUSB <> WIFE": HUSB @I13679@ -> WIFE @I13679@  [ Stefanie /Spandl/ ]
 -> "HUSB <> WIFE": WIFE @I16067@ -> HUSB @I16067@  [ Waltraut /Kratschmmar/ ]
 -> "HUSB <> WIFE": WIFE @I19672@ -> HUSB @I19672@  [ Josef /Retsch/ ]
 -> "Change SEX of HUSB": @I23247@: U -> M  [ Gerhard /Beichl/ ]
 -> "Change SEX of HUSB": @I6870@: U -> M  [ Nn /Chlup/ ]
 -> "HUSB <> WIFE": H @I566@ + W @I525@ -> H @I525@ + W @I566@  [ Gotthard /Hofrichter/ ] + [ Maria
/Rest/ ]
 -> "HUSB <> WIFE": H @I14154@ + W @I15123@ -> H @I15123@ + W @I14154@  [ Walter /Henhapl/ ] + [
Annemarie /Mast/ ]
 -> "Change SEX of WIFE": @I3284@: M -> F  [ Maria /Hlinetzky/ ]
 -> "Change SEX of WIFE": @I4787@: M -> F  [ Theresia /Fibich/ ]
 -> "Change SEX of HUSB": @I9026@: F -> M  [ Vinzenz /Bauer/ ]
 -> "Change SEX of HUSB": @I9467@: U -> M  [ Alfred /Schmidt/ ]
...
 -> "Change SEX of HUSB": @I12645@: F -> M  [ Josef /Bös/ ]
 -> "Change SEX of WIFE": @I13700@: M -> F  [ Herta /Schneider/ ]
 -> "Change SEX of WIFE": @I17107@: U -> F  [ Elfriede /Wölber/ ]
 -> "HUSB <> WIFE": H @I24997@ + W @I23522@ -> H @I23522@ + W @I24997@  [ Helmut /Neu/ ] + [ Theresia
/Kaipl/ ]
 -> "Change SEX of HUSB": @I23466@: F -> M  [ Walter /Stein/ ]
 -> "Change SEX of WIFE": @I23814@: M -> F  [ Maria /Mayer/ ]
480.254 Total lines written to modified file

Still open: 14
04.05.2015 16:00:21: Processing completed
```

The upper part informs about the ged-file and the discrepancies found.
The middle part lists the actions performed and will help during a checking at a later time.
The lower part informs about the modified file and the still open discrepancies. Note: "Open" means the not yet viewed discrepancies. Skipped discrepancies without action will be counted as "done". Thus perform always an additional run.

Note:

During reading the data, there is **no** checking of correct structures and GEDCOM syntax, valid data and consistency of the data of the ged-file. This should previously be assured by the other GSP programs.

Versions

(N = New, C = corrected, M = modified)

- 1.0.0 - 15.05.2015 - N: Erstveröffentlichung / Initial publication
- 1.0.1 - 25.02.2017 - M: Datei öffnen standardisiert / File open standardized
- 1.0.2 - 04.03.2017 - K: Datei öffnen / File open
- 1.0.3 - 01.07.2017 - M: Lizenzfunktion / License function
- 1.0.4 - 19.03.2018 - N: GedShow aufgenommen / GedShow included
- 1.0.5 - 10.05.2018 - N: "?" Show Docu