Discovering repetition of values of independent variables in EXFOR data library

V.Zerkin, IAEA-NDS, 25 February, 2014.

2013: ***A71*** *V.Zerkin   Distribute a list of data sets where repetition of value is detected in an independent variable*

Some datasets in EXFOR have full coincidence of values of all independent variables of different experimental points, although normally it should not happen [1]. The reason for that could be: mistake in compilation, precision of digitizing and low quality of figures, raw data given by authors, etc. Such repetition is allowed for data obtained in different experimental conditions and should be marked in EXFOR by flags: FLAG, MISC, MONIT*. The goal of A71 is to find datasets in EXFOR having repetitions in values of independent variables without any given explanation (potential mistakes).

Algorithm

The program works for Quantities: CS*, DA*, DE*, Ratios; it defines full coincidence of values of independent variables (including so-called "additional information", e.g. THICKNESS) and defines flag "Problem" as following:

> 0 – No coincidence
> 1 – Coincidence found, reaction code contains code RAW
> 2 – Coincidence found, DATA section has Header(s): FLAG, MISC, MONIT*
> 3 – Coincidence found

Flag "Problem" is stored in EXFOR database and used by Web retrieval system for display coinciding data since 2012.

Results

4176 datasets in 1744 Entries (i.e. 3.53% from total 118221 analysed datasets) have flag "Problem"=3: in Area 1: 448, 2:128, 3:20, 4:6, A:135, B:25, C:260, D:141, E:65, F:233, etc. Information for further analysis by EXFOR compilers available on NDS compilers site under https://www-nds.iaea.org/exfor-master/working/Zerkin2NRDC2014/A71/ in two files: list of datasets and details of comparison (positions and values of coinciding points).

Example of resulting files:
1) Listing for Area-4 (file data-repetition.out.txt)

```
===Area===[4]
597) Entry=40294: 40294005[CS]Pt:6,R:2
598) Entry=40463: 40463007[CSP]Pt:54,R:1
599) Entry=40797: 40797002[CS]Pt:87,R:2 …40797010[CS]Pt:101,R:2
600) Entry=41234: 41234002[Ratio]Pt:143,R:5
601) Entry=41369: 41369017[CS]Pt:602,R:4
602) Entry=41587: 41587003[CS]Pt:9,R:8
```

2) Details of coinciding points for SEBENT: 40797.010 (file data-repetition.tt.txt)

```
#SubentP:40797010  Q:[CS] R:51-SB-0(N,G),,SIG
 NumberWithSameXX:2 iProblem:3  DataRows:101
 X-Vars: EN(MEV);
#DATA 2*101
   80)   81) 0.0045;
   99)  100) 0.022;
#ENDDATA 2*2
```

References

1. Indication in EXFOR database and Web interface the full coincidence of values of independent variables of different data points in one dataset. V.Zerkin, WP2013-34, NRDC2013, 23-25 April 2013