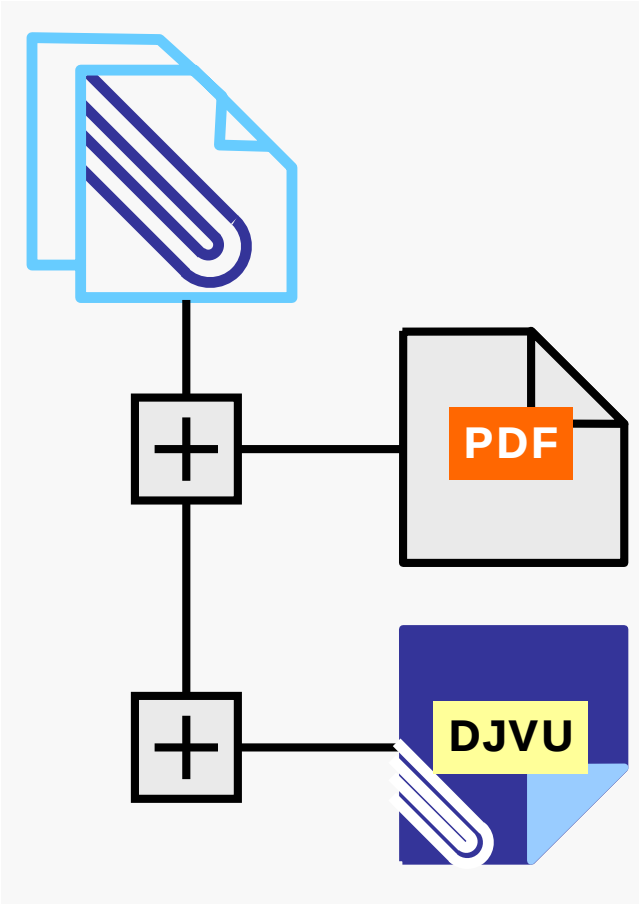
Andrew Stolyarov's

Pdf & DjVu Bookmarker™ v. 3.5.9

User manual v.3.0.4



NBell

2012

About User manual

User manual is written by Nbell.

Download page on RUTRACKER

License and copyright

The rights on design and content of this User manual belongs to NBell. Any usage of this User manual or parts of it permitted with reference on the NBell.

User manual provided «as is» without any warranty.

©NBell 2012

No animal suffered from this manual!

Collaboration

The author suggests his help in creating technical documentation on Russian language.

Contact with NBell via e-mail **bellick-n(dog char)yandex.ru**

Adobe, the Adobe logo, Acrobat, the Adobe PDF logo and Reader are either registered trademarks or trademarks of Adobe Systems Incorporated in the United States and/or other countries.

Caminova, Document Express are trademarks and DjVu is a registered trademark of Caminova Inc. in the United States and/or other countries.

Microsoft, Windows, .NET, Word are either registered trademarks or trademarks of Microsoft Corporation in the U.S.and/or other countries.

WinDjView Copyright (C) 2004-2009 Andrew Zhezherun

DjVuLibre project maintained by Leon Bottou, Yann LeCun, Bill C Riemers, and Jeffery Triggs

Table of content

About User manual 2
Table of content 3
1. About 5
1.1 License 5
1.2 System requirements 5
2. What is «Tree Of Content (TOC)»? 6
3. Window «TOC Source Selection» 7
4. Menu «File» and menu «Tools» 8
5. Toolbar «Preview» 10
6. Num Editor 11
6.1 Window and commands 11
6.2 Toolbar «Check» (Num check) 12
6.3 Modes of «Errors» window 12
6.4 Toolbar «Merge» 12
6.5 Toolbar «Replace» 14
6.6 Window «Leadering remove» 16
6.7 Structured TOC creation 17
7. Structure Editor 18
7.1 Window and commands 18
7.2 TOC level preview 18
7.3 Toolbar «Edit numbering» 18
7.4 Toolbar «Edit» 19
7.5 Toolbar «Structure» 20
7.5.1 Group shift 20
7.5.2 «by template» 21
7.6 Toolbar «Position» and Menu «Apply position» (Position Editor) 22
8. Saving TOC 24
9. Tips and tricks 25
9.1 Quick learning TOC creation 25
9.2 Spellcheck 25
9.3 Copy TOC from one e-book to another 25
9.4 Erase TOC 25
9.5 Extract TOC 25
9.6 Remove structure from TOC 25
9.7 How to get TOC text from e-book 25

10. Known problems 27
10.1 MS .NET 2.0 27
10.2 Error with Acrobat as default pdf viewer 27
10.3 Error «exceeded the number of pages» 27
10.4 Error «Unhandled exception» 27
10.5 DjVu file does not opened for preview or editing TOC 28
10.6 Unexpected end of file 28
10.7 This node contains more then 255 subnodes 29
11. Appendix 30
11.1 TOC file format (txt, html) 30

1. About

Pdf & DjVu Bookmarker can help you to create tree of content (Outline, Bookmarks) in djvu and pdf e-books.

Main features:

- WYSIWYG editing and creating tree of content (TOC).
- Smart TOC consistency check.
- Special command set for editing text after OCR.
- Save TOC in txt or html file format (UTF-8 encoding).

Creator of Pdf & DjVu Bookmarker – Stolyarov Andrew

Download page on SourceForge net

Discussion topic on ru-board

The program uses for preview of

- DjVu: Djvu Libre (with no installed Caminova Document Express DjVu plug-in);
- PDF: PdfLibNET (with no installed Adobe Reader).

1.1 License

GNU GENERAL PUBLIC LICENSE Version 3, 29 June 2007

1.2 System requirements

See requirements of .NET 2.0.

Required additional software:

Microsoft .NET Framework 2.0 (x86) 22.4M6

Recommended software for confortable viewing:

Caminova Document Express DjVu plug-in;

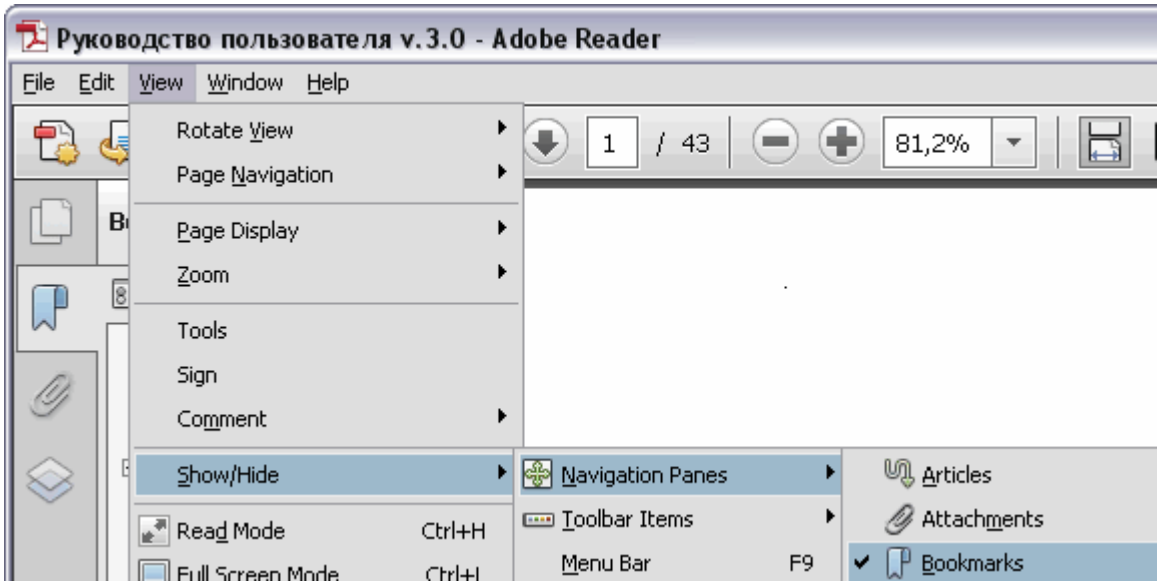
Adobe Reader X.

Note: If you have installed Acrobat/Acrobat Pro as default pdf viewer program raise error message on exit. Install Adobe Reader with option «Make Adobe Reader my default viewer»(or if Adobe Reader already installed, choose Adobe Reader as default pdf handler via «Edit – Preferences – General – Select Default PDF Handler - Select Adone Product» and select «Adobe Reader»).

2. What is «Tree Of Content (TOC)»?

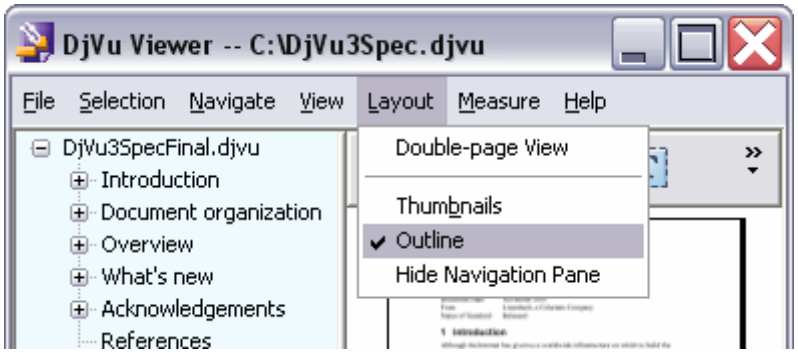
Easy to use e-book must have a tree of content for navigation.

In Adobe Reader - Bookmarks.

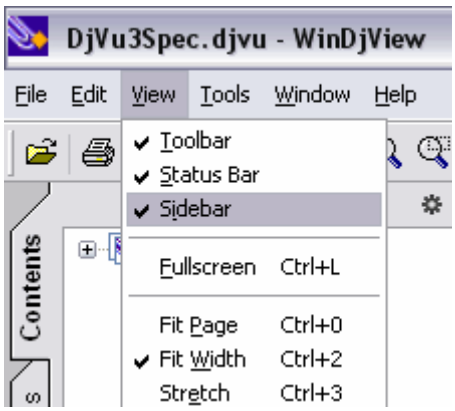


In LizardTech Djvu Browser plug-in (Caminova Document Express DjVu plug-in)

«Layout - Outline»

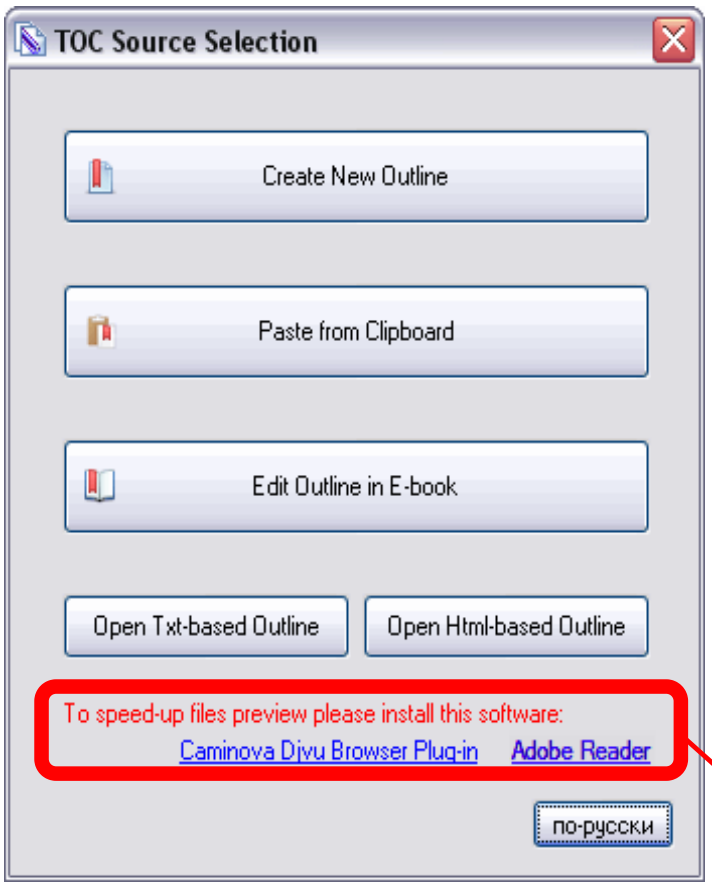


WinDjView TOC appears on «Sidebar» panel (tab «Contents»).



DjView uses tab «Outline» on sidebar.

3. Window «TOC Source Selection»



Program checks availability of ***DjVu plug-in*** and ***Adobe Reader***. This message appears if this viewers not installed. Blue links are clickable and opens default browser with download page.

«Create New Outline» - program create blank TOC and go to the num editor.

«Paste from Clipboard» - gets Clipboard content and go to the num editor.

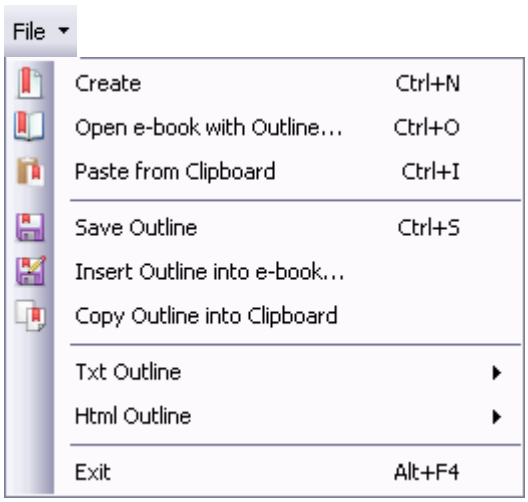
«Edit Outline in E-Book» – open PDF or DJVU e-book and go to the TOC editor. (Some e-books may have no or non-standard TOC).

«Open Txt-based Outline» - get TOC from txt-file and go to the num editor.

«Open Html-based Outline» – get TOC from html-file in BookmarkTool 2.0-format. and go to the TOC editor. (Some kind of links not supported).

«по русски» - change program language to «Russian».

4. Menu «File» and menu «Tools»



«Create New Outline» (Ctrl+N) - program create blank TOC and go to the num editor.

«Open e-book with Outline» (Ctrl+O) – open PDF or DJVU e-book and go to the structure editor. (Some e-books may have no or non-standard TOC).

«Paste from Clipboard» (Ctrl+I) - gets Clipboard content and go to the num editor.

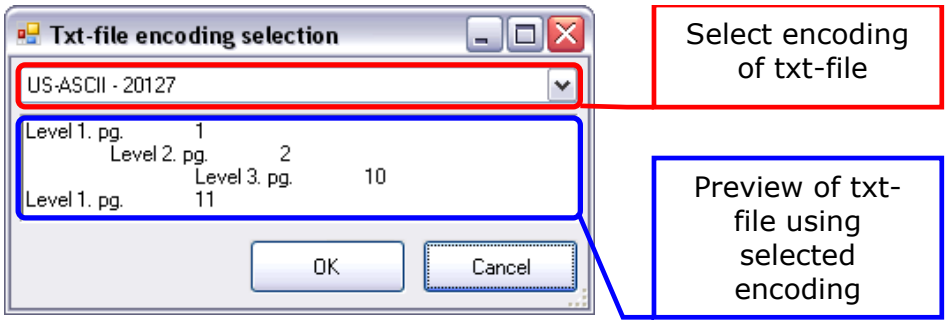
«Save Outline» (Ctrl+S) – embeds created or edited TOC in the opened e-book.

«Insert Outline into e-book» – embeds created or edited TOC in the selected e-book.

«Copy Outline into Clipboard» – copy created TOC to clipboard.

Txt Outline – submenu for open/save TOC in txt file format.

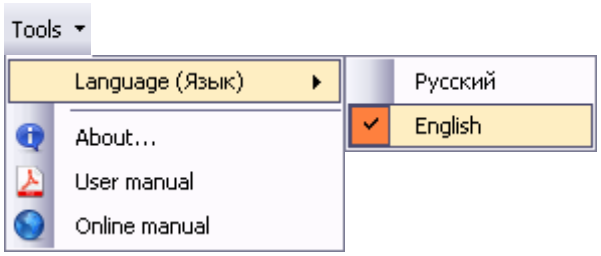
After txt-file opening encoding can be WYSIWYG - adjusted:



Html Outline – submenu for open/save TOC in html (BookmarkTool 2.0) file format. This command available in structure editor.

Exit (Alt+F4) – close program.

Use «Tools» menu for change interface language.



About… – shows program’s credits.

User manual – opens this User Manual in the default pdf viewer.

Online manual – opens wiki-documentation in the default Internet browser.

5. Toolbar «Preview»

Use this toolbar for preview of e-book in «Preview window». Supported djvu and pdf formats.

«select e-book» - for choosing file for view.

«show pages» - toggle-button appears if opened any file for preview. Turns on/off page displaing in the preview window (reduces CPU usage on the slow PC).

Navigation blue arrows with textbox between appears and used for navigation in built-in preview if no Adobe Reader and/or DjVu plug-in installed. For preview previous / next page you can use hotkeys Ctrl+PgUp / Ctrl+PgDown.

Adobe Reader and DjVu plug-in has their own navigation tools.

Adobe Reader - use F8 to show navigation toolbox.

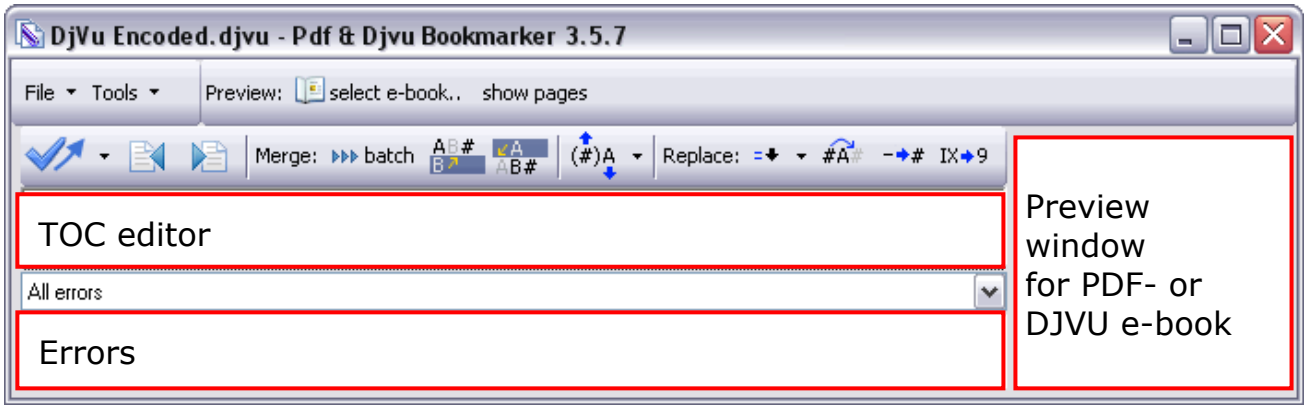
DjVu plug-in - shows toolbars when mouse pointer moves to top/bottom of preview window (depends on your DjVu plug-in settings). Deselect «Auto Hide» in the context menu of toolbar for constant toolbar appear. Docking place depends on setting «Top» or «Bottom».

In structure editor preview-only mode available when «Edit Numbering» turned off.

6. Num Editor

This editor used when new TOC created, pasted from Clipboard or txt-file opened.

6.1 Window and commands



Inner windows border adjustable – drag it with left mouse button pressed.

«TOC editor» - simple text editor of created TOC. Standard hotkeys:

- Select all Ctrl+A;
- Copy to clipboard Ctrl+C;
- Paste from clipboard Ctrl+V;
- Cut to clipboard Ctrl+X;
- Undo Ctrl+Z;
- Line end – Enter.

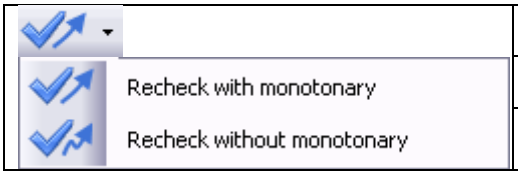
Text size can be adjusted by Ctrl-mouse wheel. Edit cursor must be in window which text size adjusted (simply click in any line).

Combo box below TOC editor is filter for errors. See next chapters.

«Errors» show TOC errors. Click on line shows the TOC-line with error.

«Preview» window show opened e-book.

6.2 Toolbar «Check» (Num check)

	«Recheck». Check TOC errors (F7). Check all errors Exclude monothony check

shows previous error (Alt + ↑, F5).

shows next error (Alt + ↓, F6).

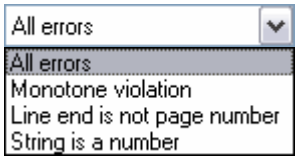
Program allows TOC (new, pasted from clipboard or imported from txt-file) to be embedded in e-book only after consistency check (Num check).

Num check starting by pressing «Recheck» (F7) button. List of errors shown by «Errors» window.

Error message	How to correct
Monotone increasing of page numbers violation	Numeration of previous line must be less or equal checked line. Correct number or, if you advanced user, skip this check by choosing «Recheck without monotonary» command. See above.
Whole string is a number	Line has no text, only number. Remove the line or enter some text you need.
Line end is not page number	Line ends with non-mumeric character. Enter number, or correct error with commands, described in next chapters.

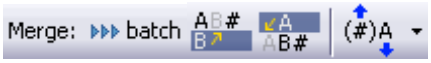
6.3 Modes of «Errors» window

On the top of «Errors» window located combo box with options:



Note: Background of «Errors» window changes to green when errors filtered.

6.4 Toolbar «Merge»



Useful when lines splitted after OCR or merged (by OCR or text design).

merges line without number with previous line, ends with number (Ctrl + ↑, F1).

Example

Source	After
Chapter 1. Rules of ……1 the game	**Chapter 1. Rules of …… the game 1**

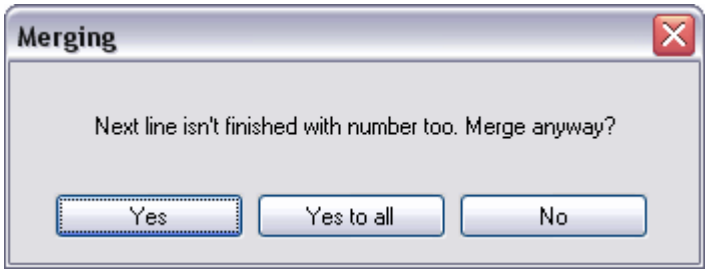
merges line without number with next line, ends with number (Ctrl + ↓, F2)

Example

Source	After
Chapter 1. Rules of the game ……1	**Chapter 1. Rules of the game ……1**

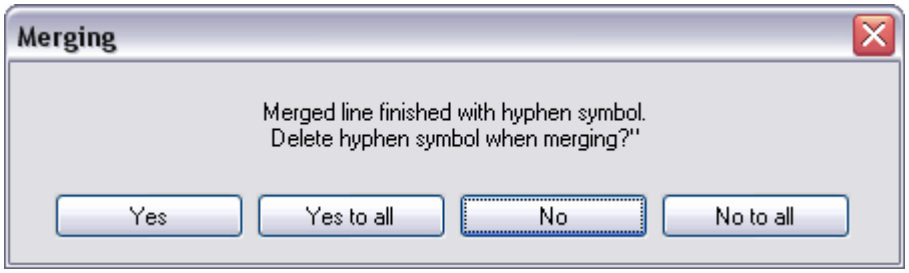
toggle-button turns on/off batch merging.

If merging line contains no number you get message



Click «Yes» to merge anyway. «No» – to cancel merging.

If line ends with hyphen «-», program ask what to do:



Click «Yes» to remove hyphen and merge. «No» – merge with hyphen.

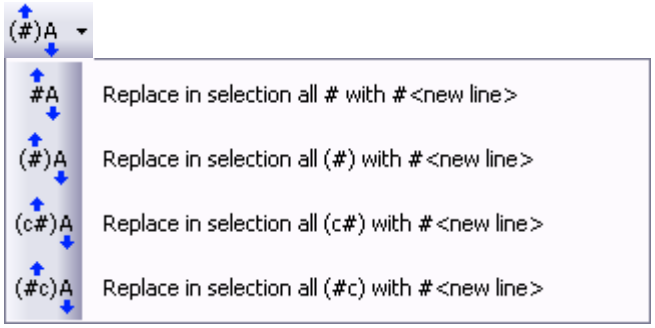
splits lines of TOC. Sometimes after OCR or by design TOC looks such:

Example:

Chapter 1. Rules of the game (1) Chapter 2. Play (3)

Chapter 1. Rules of the game…..1Chapter 2. Play……..3

Select merged line and split, using suitable command:



«#» - means the page number, «A» – text line, «c» – some text in brackets near page number. If cursor in line, but line not selected, click twice to apply the command.

- simple line splitting merged after OCR - «Chap1…3Chap2…5»;

(#)A - splits «Chap1 (3) Chap2 (5)»;

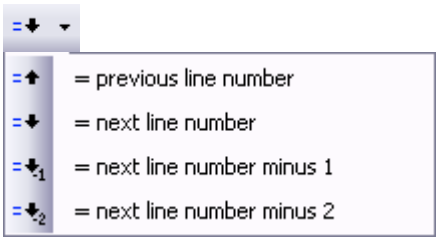
(c#)A - splits «Chap1 (c. 3) Chap2 (c. 5)»;

(#c)A - splits «Chap1 (3 c.) Chap2 (5 c.)».

6.5 Toolbar «Replace»

Useful for fast numbering of lines.

add page number to line without it (F9). Apply to selection.



Example:

Source:	Command	After:
Chapter 1. Main 5 **1. Value……**		**Chapter 1. Main 5** **1. Value……5**
Chapter 1. Main **1. Value……5**		**Chapter 1. Main 5** **1. Value……5**
Chapter 1. Main **1. Value……5**		**Chapter 1. Main 4** **1. Value……5**
Chapter 1. Main **1. Value……5**		**Chapter 1. Main 3** **1. Value……5**

#A moves page number from the beginning of the line to the end of.

Example:

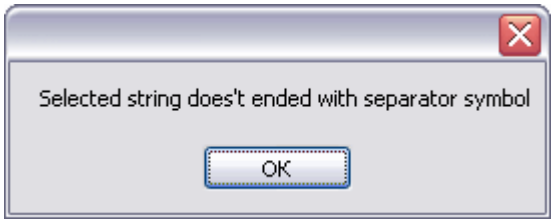
Source:	After #A
5 Intro **6 Foreword**	**Intro 5** **Foreword 6**

-># replaces all substitute symbols to nearest upper page number. Substitute symbol – any non-numeric and non-text character, located at the line end. Be careful - sometimes OCR produces lines with page number and non-numeric symbols at the end of line.

Example:

Source:	After ->#
Chapter 1 44 **Chapter 2 -** **Chapter 3 -** **Chapter 4 45** **Chapter 5 - -**	**Chapter 1 44** **Chapter 2 44** **Chapter 3 -** **Chapter 4 45** **Chapter 5 - 45**

To use command move the cursor to the line with substitute symbol («Chapter 2», «Chapter 3» or «Chapter 5»). If the cursor on line with number at the end («Chapter 1»), then after command you get error:



«Chapter 3» needs to be selected and command must be reapplied.

IX→9 replaces Roman numerals to Arabic numerals.

As usual, Roman numerals used at the start of the book. And page numbers you get will be greater, than latter Arabic. Use «Recheck without monotonary» and correct numeration later in structure editor.

Example:

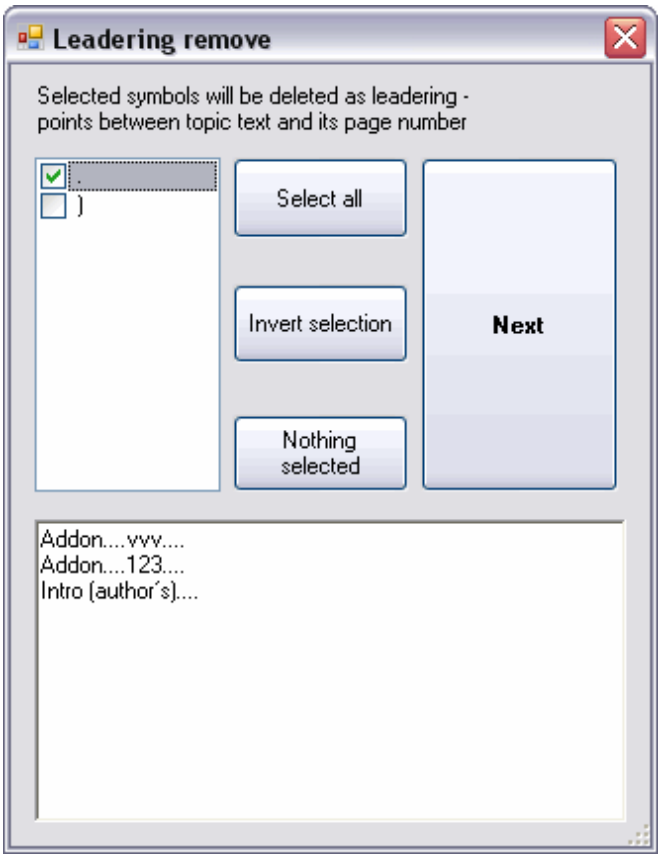
Source:	After IX→9:
Intro V **ever** **be VII** **exist 10** **later XX**	**Intro 5** **ever** **be 7** **exist 10** **later XX**

Note: Command replaces Roman numerals up to first line with Arabic (line «exist 10»).

6.6 Window «Leadering remove»

With no leadering (any symbol between last character of text=A...Z, a...z, A...Я, a...я and page number at the line end=0...9) this window doesn't appear.

After success in num check and, if the text contains any leadering symbols between text and page number, you get:



Select all leadering to remove «**Select all**».

Change selected to unselected «**Invert selection**».

Select nothing (all symbols considered as useful) «**Nothing selected**».

Remove leadering by button «**Next**».

When leadering in upper listbox selected, lower listbox show lines, containing it.

Example:

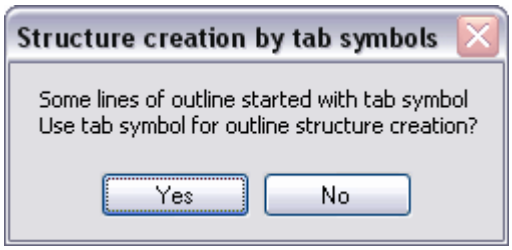
Source:	After «Next» as shown:
Addon....vvv....1 **Addon....123....1** **Intro (author's)....100**	**Addon....vvv 1** **Addon....123 1** **Intro (author's) 100**

Flag near «)» cleared – it is the part of the text.

OCR errors "....vvv" and "....123" had to be deleted manually in structure editor.

6.7 Structured TOC creation

If the TOC contains tab symbols (see TOC file format):



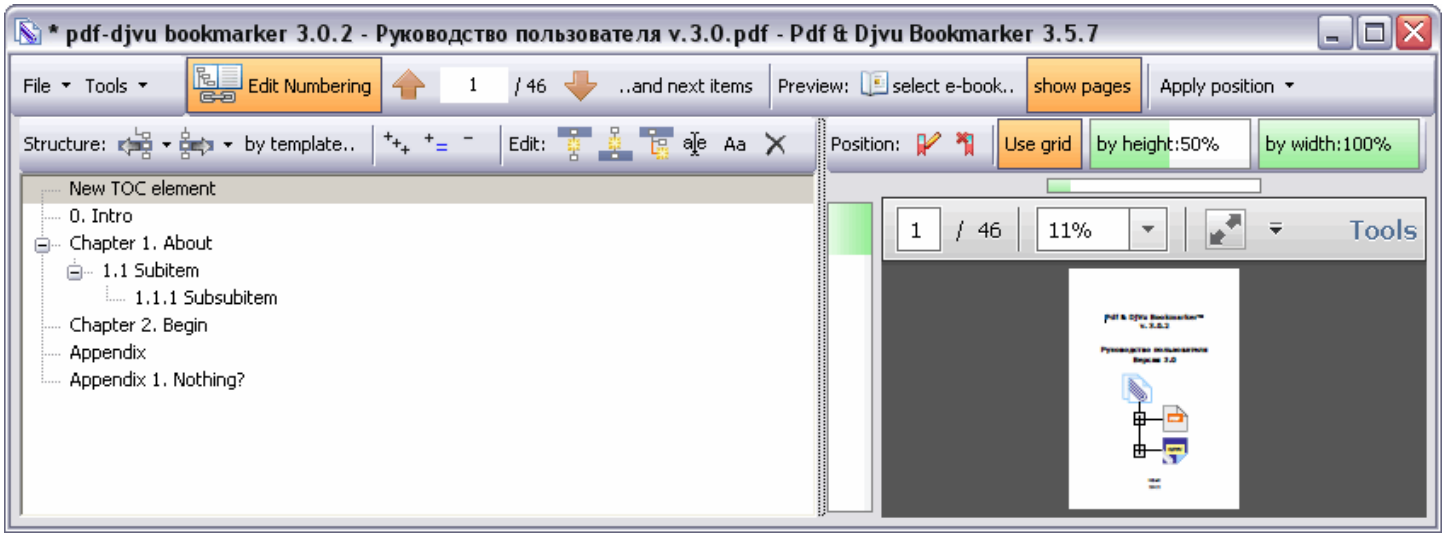
Press «Yes» if you want to set structure by tab symbols.

If no structure needed - press «No».

7. Structure Editor

Useful if multilevel TOC structure needed. Also you can edit numbers and text.

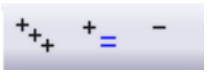
7.1 Window and commands



Left window shows the TOC - «TOC window», right shows e-book – «Preview». File to preview choose by «Preview» toolbar. Selection of the TOC line can be changed by cursor keys, corresponding page immediately shown in «Preview» window after pressing «Enter».

Note: Use «Ctrl+A» hotkey to select all TOC elements.

7.2 TOC level preview



Adjust view of the multilevel TOC by:

- expands all levels;

- expands levels with rank equal to selected.

- collapses TOC items to top-level.

7.3 Toolbar «Edit numbering»

Toggle-button «Edit Numbering» turns on/off TOC item page number value editing. If toggle-button «..and next items» pressed – all links of TOC items below selected changes too (relative to changed item).

WYSIWYG mode of TOC links adjusting realized if appropriate e-book opened in «Preview».

To change link in selected item only depress «..and next items» (no blue plus below arrows):

Depressed button «Edit Numbering» allows browsing e-book in «Preview» when built-in preview used.

7.4 Toolbar «Edit»

Useful to edit items text, adding new items, deleting items.

adds new item below selected (Insert).

adds new item above selected (Shift+Insert).

adds new child to selected item (Ctrl+Insert).

edits item text (F2).

removes item (Delete). Note – all child items, if any exists, will be deleted too.

changes case similar to sentence (Ctrl + Alt + O). Remember:

- Accurate works with Roman numerals (inaccurate, if text contains Roman numerals only).
- After any numeral makes next word with capital letter.
- Word after the dot starts with capital letter.
- First word always starts with capital letter.
- If needed capital letter in other word – correct manually.

Example:

Source:	After :
CHAPTER 1 START **CHAPTER V START** **CHAPTER 1. START** **CHAPTER ONE. START** **CHAPTER ONE. PETER PAN**	**Chapter 1 Start** **Chapter V Start** **Chapter 1. Start** **Chapter one. Start** **Chapter one. Peter pan**

Word «pan» must be manually corrected by .

7.5 Toolbar «Structure»

Useful in changing level of the TOC item.

increases level, Alt+ ←, mouse gesture – hold RMB then LMB.

decreases level, Alt+→, mouse gesture – hold LMB then RMB.

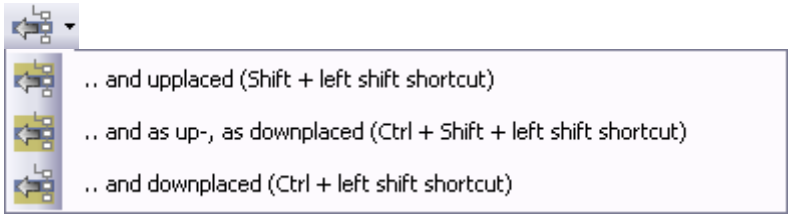
Group select with Shift or Ctrl similar to Windows Explorer.

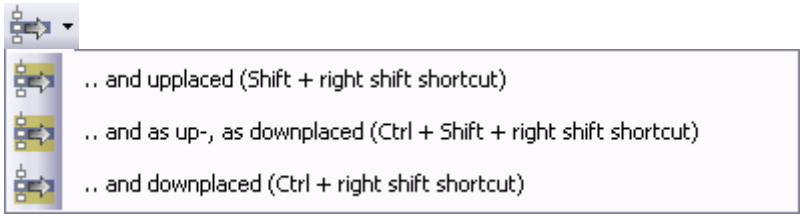
7.5.1 Group shift

You can select one TOC item and use structure command with:

- Ctrl – shift selected and below
- Shift - shift selected and above
- Ctrl+Shift - shift all.

Modified commands available from dropdown menu:



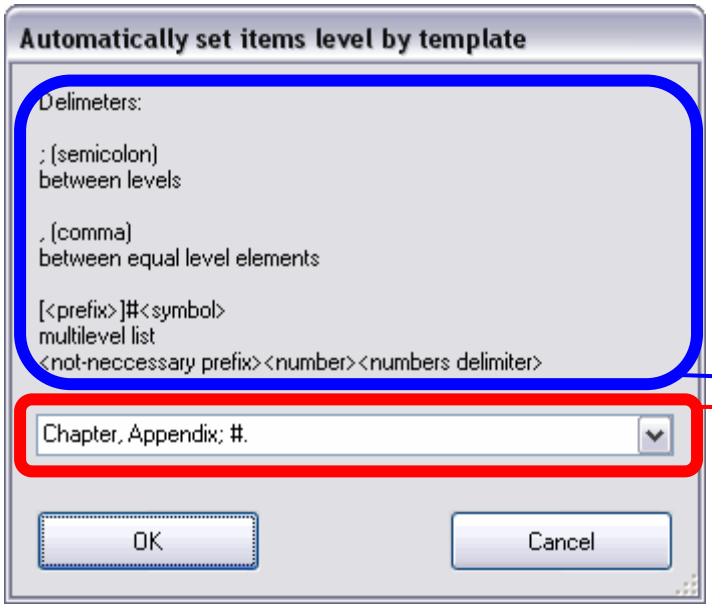


For left group shift command will be applied to child items too.

For right group shift command will be applied to childless items only.

7.5.2 «by template»

If source TOC have logical structure, then you can use it for creating structured TOC. Press «by template». Appears window:



Rules of templating:
Every new level separated by «;».
Items within level – by «,».
Numbered list set by
[prefix (any repeating word or symbol before page number)]
#(list item number)
[symbol after number (usually dot or some other symbol)]

In this combo box you can select one of the built-in templates. You can enter any valid custom template. See rules above.

Example:

#.

-<number>, <symbol> - the dot. Items with names 1, 1.2, 1.2.4 и т.п. (no prefix used) will be treated as multilevel list and «1» get level 1, «1.2» - 2 and «1.2.4» - 3.

Example:

P#.

<prefix> - symbol «P», # -<number>, <symbol> - the dot. Items with names P1, P1.2, P1.2.4 will be treated as multilevel list and «P1» get level 1, «P1.2» - 2 and «P1.2.4» - 3.

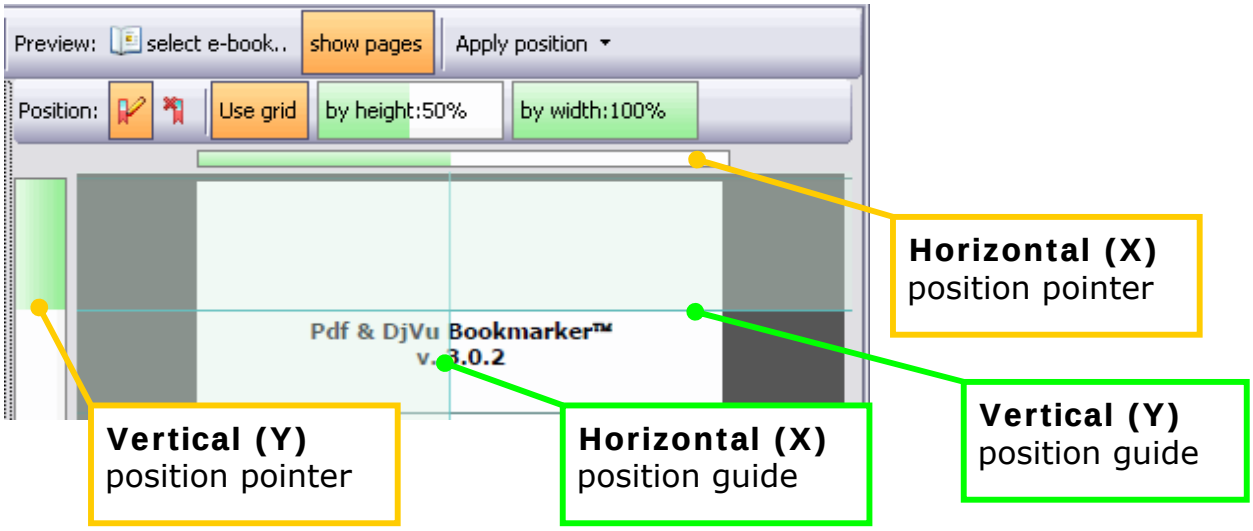
Example below means, that items, started with «Chapter» or «Appendix» will be top-level, all other below (items, begins with <number><dot> get level 2 and lower in order of multilevel list, but only if parent item exists – see «0. Intro»).

Example:

Source:	By template «Chapter, Appendix; #.»:
0. Intro **Chapter 1. About** **1.1 Subitem** **1.1.1 Subsubitem** **Chapter 2. Begin** **Appendix** **Appendix 1. Nothing?**	0. Intro Chapter 1. About 1.1 Subitem 1.1.1 Subsubitem Chapter 2. Begin Appendix Appendix 1. Nothing?

7.6 Toolbar «Position» and Menu «Apply position» (Position Editor)

Useful (in pdf e-book) in setting the position on page, showed after TOC item selected.



In pdf e-book TOC item show upper left page corner by default. Click LMB on position pointer to change show position (click Y-pointer – set vertical, H-pointer - horizontal).

Note: All TOC items property «Zoom level» will be set to «Inherit zoom».

toggle-button turns on/off position editor.

Note: Fast set of Y-position can be done by selecting TOC item and pressing «Enter». Set position by cursor keys «↑» and «↓».

Attention!!! Hide Adobe Reader toolbar when using position editor! (Click on preview window and press F8 if Adobe Reader toolbar not hidden.)

Grid is drawn by program to make position setting easier.

- toggle-button turns on/off grid.

rulers, which set the grid frequency.

«by height» – accuracy of seek free space between text lines.

«by width» – accuracy of seek free space between text and vertical page border.

Maximum frequency of the grid corresponds to 0% «by height» and 100% - «by width».

Aplly position by:

- LMB click in desired place on page;
- press «Enter».

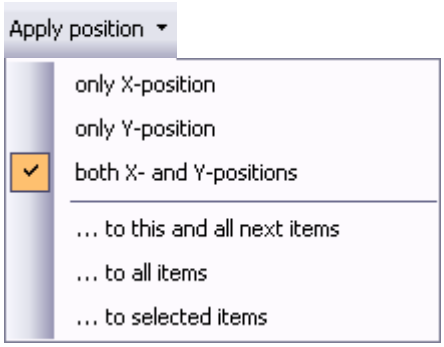
Exit from position editor without changes by:

- depressing .
- RMB click on page;
- LMB click on one of position pointers;
- press «Esc».

Note: Page scale in the preview window must be set to see whole page.

sets position to default (upper left page corner).

TOC item show position can be applied to other items by menu:



Selected TOC item position using to:

- set all horizontal positions to current «only X-position»;
- set all vertical positions to current «only Y-position»;
- set all horizontal and vertical positions to current «both X- and Y-positions».

Methods of apply:

- from selected and below to the end – «…to this and all next»;
- to all TOC – «…to all items»;
- for selected items – «…to selected items».

Note: Position correctly shown only in Adobe Reader (or other viewer). It's a pity, but you can't see it in WYSIWYG mode when editing – only by position pointers.

8. Saving TOC

In Num editor TOC can be saved in txt-file only. Use «File – Txt Outline – Save Txt As».

In Structure editor command:

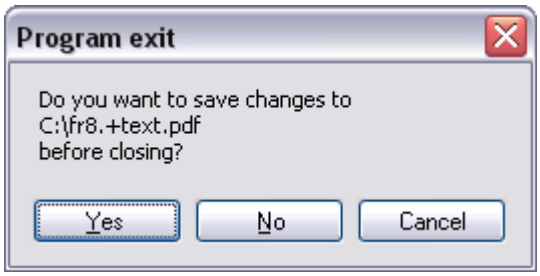
«File – Save Outline» (Ctrl+S) embeds created structured TOC in e-book, opened in «Preview» window. Inactive when no e-book selected for preview.

«File – Insert Outline into e-book» embeds created structured TOC in e-book, selected in standard dialogue window.

«File – Txt Outline – Save Txt As» - saves created structured TOC in txt-file.

«File – Html Outline – Save Html As» - saves created structured TOC in html-file.

If you try to exit program in structure editor without saving TOC appears message box:



9. Tips and tricks

9.1 Quick learning TOC creation

Read this manual. Examples source text can be copied to clipboard and pasted in TOC editor using Ctrl+I.

Use hotkeys. Read the tooltips.

9.2 Spellcheck

Copy created TOC to clipboard and paste in text editor with spellcheck. Correct errors, copy the TOC to clipboard and paste in program.

9.3 Copy TOC from one e-book to another

Open e-book with TOC «File – Open e-book with Outline». Embed TOC to another file «File - Insert Outline into e-book».

9.4 Erase TOC

Sometimes wrong TOC causes errors. Create new TOC with any item, for example «TOC 1», and embed it into the e-book with wrong TOC.

You can use DJVUSED (shipped with DjVuLibre).

djvused –e "remove-outline" djvufile -s

9.5 Extract TOC

If TOC text needed for some purpose, open e-book «File – Open e-book with Outline». Copy TOC to clipboard and paste in text editor or save in txt or html file.

9.6 Remove structure from TOC

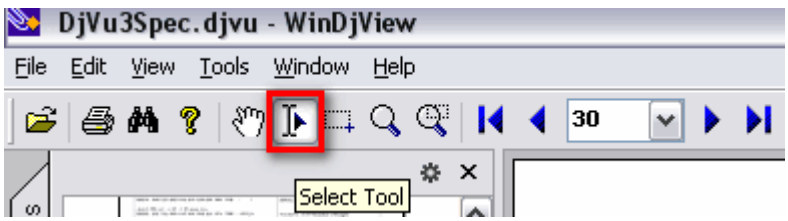
Fast remove of undesired structure:

1. In structure editor copy TOC to clipboard using «File – Copy Outline into Clipboard».
2. «Paste from Clipboard».
3. Recheck numeration (F7).
4. In message window «Structure creation by tab symbols» press «No».

9.7 How to get TOC text from e-book

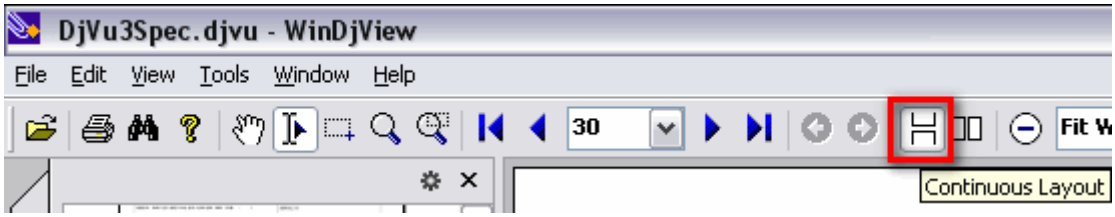
DJVU

Open file in WinDjView. Choose «Select Tool».



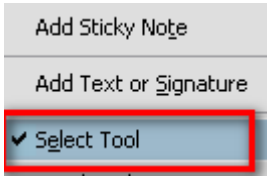
Select table of contents text, copy to clipboard and paste to program.

If text placed on two and more pages choose «Continuos Layout» before selecting:



PDF

In Adobe Reader click RMB on text and choose «Select Tool»:



If text placed on two and more pages choose «View – Page Display – Enable Scrolling» before selecting:

If no text layer present, copy images to OCR-program.

Sometimes table of content can be founded by Internet search.

10. Known problems

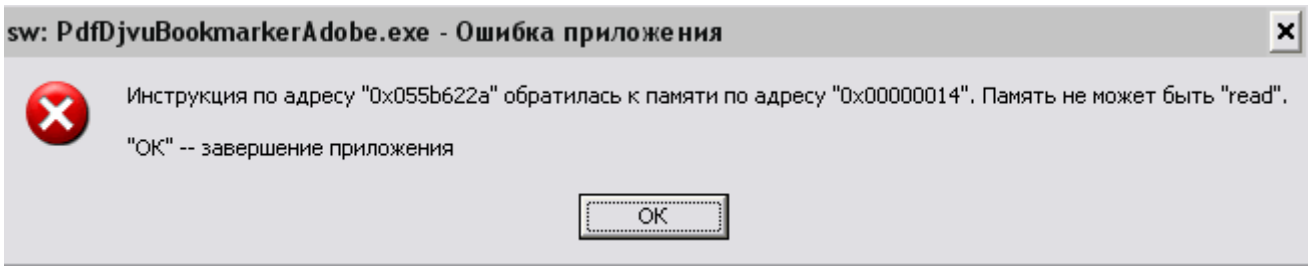
10.1 MS .NET 2.0

If MS .NET 2.0 not installed, program show error message

See «System requirements» for MS .NET 2.0 download link.

10.2 Error with Acrobat as default pdf viewer

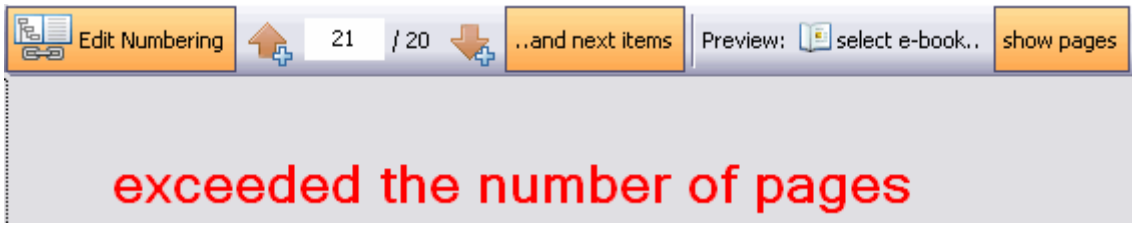
After pdf e-book open for preview and exit error message may appear:



This error does not affect on edited e-book. You can use recommendations.

10.3 Error «exceeded the number of pages»

Appears if page number link setted (21 in textbox) bigger than e-book page count (20 after slash), or setted to zero (0<1 – minimum page count).



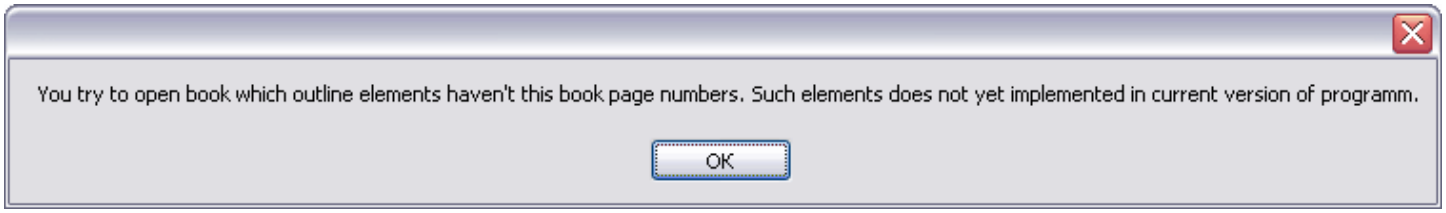
Sometimes appears with correct links. Save TOC – e-book will not be damaged.

10.4 Error «Unhandled exception»

It appears if wrong parameters set. Click «Continue» and continue work.

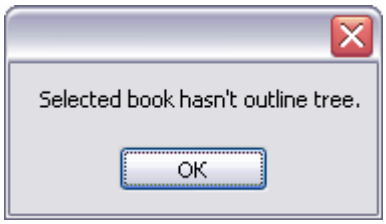
10.5 DjVu file does not opened for preview or editing TOC

Program supports djvu TOC with item link type «Page Number» only. If TOC contains blank links, named links («Page Name») or URLs, you get message:



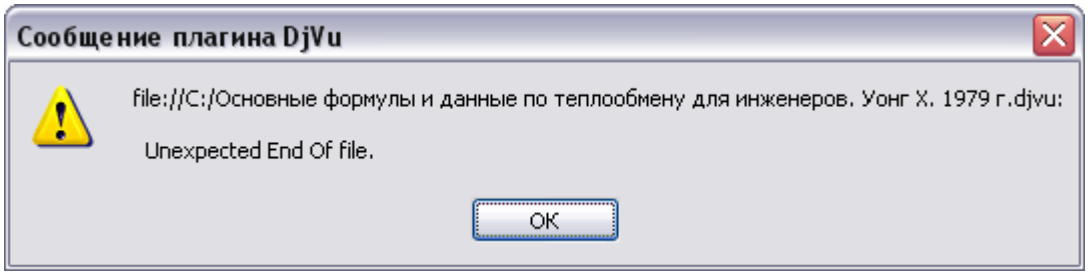
Such TOC can be edited in Document Express Editor or other editor.

Some pdf e-books may contain non-standard TOC (with page names, javascript, etc.) or no TOC at all. In that case when you try to open it for editing TOC appears message:



10.6 Unexpected end of file

If you you choose standard Caminova viewer and open djvu with non-standard TOC (but Windjview or DjView opens the file without errors) appears message:



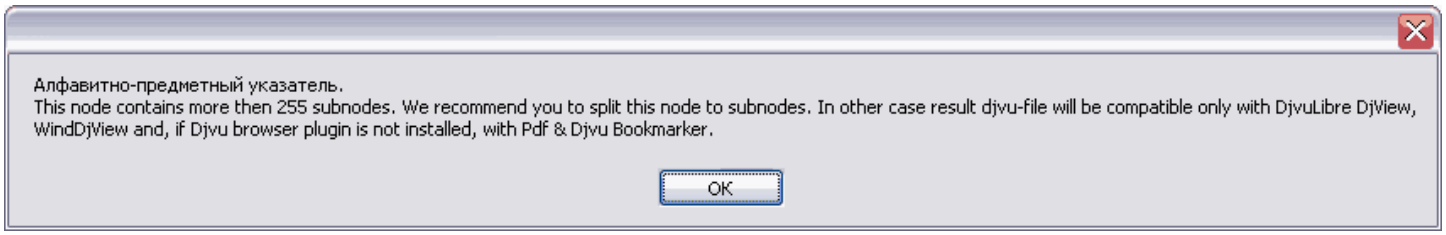
Error message appears if top-level item count more then 255 or if one or more child branches contains more then 255 items. Djvu specification says, that the count of child items of one parent item is “byte” (0-255). So, child number 256 or larger is illegal/

Split long item by creating two or more chid items and move there parts of the long item with count <256.

Long TOC does not supported by official commercial djvu viewers and editors. Windjview, DjView can show such files without any error messages? But such file becomes nonstandard.

10.7 This node contains more then 255 subnodes

This error message appears if you try to save non-standard TOC (where one or more items have 256 or more child items):



DjVu 3 specification allow only 255 chid items for one parent item. You can split problem item on two or more subitems and place childs there (255 or less on one item).

Or, if you use nonofficial viewers (Windjview, DjView), pree “OK” and TOC will be embedded. If you try to edit it again, program (if you choose standard Caminova viewer) raise error “Unexpected end of file”.

11. Appendix

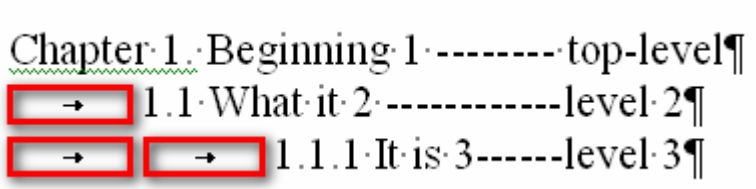
11.1 TOC file format (txt, html)

Program works with txt and html files encoded in UTF-8.

Save TOC files in UTF-8 (UTF-8 without BOM) encoding.

In txt file levels can be set with tabulation characters.

Example (MS Word, tabulation outlined by red):



Note: Tabulation characters visible if button pressed ¶ (Ctrl+*).

Use HTML-file (BookmarkTool 2.0 format) in UTF-8 encoding **(set encoding whike saving text. Tag «charset» will be ignored).**

Example:

```
<html>
<body>
<ul>
<li><a href="#1">Level 1. Link to page 1</a></li>
<li><a href="#2">Level 1. Link to page 2</a></li>
      <ul>
      <li><a href="#4"> Level 2. Link to page 4</a></li>
      <li><a href="#5"> Level 2. Link to page 5</a></li>
      </ul>
</ul>
</body>
</html>
```

Page number set in tag «a href = «#(page number)». Level can be set by tags ….

Note: Named links or file links (URLs) not supported.