

PLATYPUS FUNCTIONAL REQUIREMENTS v. 2.02

TABLE OF CONTENTS

Introduction.....	2
Input Requirements	2
Input file.....	2
Input File Processing.....	2
Commands	3
Categories of Commands	4
Formatting Commands.....	4
Page Formatting Commands	4
Paragraph Formatting Commands	5
Non-Font Text Formatting Commands	6
Font Formatting Commands.....	7
Other Document Element Formatting Commands.....	8
Action Commands	9
Special Characters	11
Languages.....	11
Special Punctuation.....	12
Mathematical AND Scientific.....	13
Miscellaneous Characters.....	14
Other Requirements	14

INTRODUCTION

Platypus reads an input text file containing text and Platypus commands from which it generates a document in a supported format. Supported formats include PDF, HTML, Microsoft Word, text, and a simple listing format that produces a colorized listing of the Platypus input file. Other output formats can be developed using plug-ins.

INPUT REQUIREMENTS

INPUT FILE

- I01 Input file is a text file
- I02 Input file may contain UTF-16 characters
- I03 Input file can include other files
- I04 Included files are read into the principal file at the point where they include command appears.
- I05 Included files can nest up to an unlimited depth
- I06 Included files cannot call each other: every included file in the chain must be unique
- I07 An input file that contains no Platypus commands should generate a valid output file. (This is made possible by extensive use of sensible defaults.)

INPUT FILE PROCESSING

- I21 Input files may contain: text, commands, and comments
- I22 Comments are saved or discarded based on a configuration option of the output format. This allows future output plugins to extend Platypus by passing new commands through comments.
- I23 Commands can contain scripts and/or script-language commands.
- I24 The specific script language [has yet to be determined]. It must be able to:
 - I24a Be passed parameters
 - I24b Access Platypus variables
 - I24c Resolve logical and arithmetic expressions
 - I24d Return a string (text or command) or a numeric value

- I24e Be easy to learn quickly [preference given to existing scripting language]
- I25 Commands can contain macros
- I26 Macros are pure string substitution
- I27 Macros can call other macros to a depth of two
- I28 Macros cannot call themselves recursively
- I29 Macros may contain commands and are rescanned for those commands at output time.
- I30 Text can be modified several ways: input plugins, command-line parameters, commands
- I31 Input plugins are specified on the command line. They are applied to text only and can generate commands. They allow text with embedded commands that are unknown to Platypus to be processed correctly. (For example the use of `_` before and after a word to convey the use of italics.)
- I32 There is no preset limit on the number of input plugins that can be used. They are executed in the order specified on the command line.
- I33 Once an input file has been completely parsed into commands, text, comments and all included files equally parsed, the resulting file is passed to the output plugin appropriate for the format.
- I34 After the input file is passed to the output plugin, the input routines can still be called, such as to parse the result of resolving a macro, which could entail a series of commands.

COMMANDS

- C01 Commands are designed to be intuitive
- C02 Frequently used commands should be short
- C03 All commands use a closed syntax. That is, the beginning and end of the command are known by looking at the command. Some commands that affect large chunks of text use a begin-command and end-command pair to designate the affected text.
- C04 Syntax must provide for commands that take parameters.
- C04a Syntax must provide for multiple parameters.
- C05 Some commands are compound commands. They represent commands that involve several parameters or actions that must go together or must be performed together.

CATEGORIES OF COMMANDS

- C10 There are two types of commands: input and output.
- C11 Input commands affect the processing of the input file. (For example, including additional input files.)
- C12 Output commands indicate formatting specifications, Platypus actions, or insertion of special characters.
- C13 Actions are generally expressed via the scripting language, although not always.
- C14 Platypus provides extensive debugging features. These features are implemented using action commands, generally dumps of variables and other pertinent information.
- C14 Formatting specifications express how the output should appear in the output document.
- C15 Special characters are characters or symbols that are difficult to express in text, such as foreign characters, mathematical symbols, etc.

FORMATTING COMMANDS

Note: Some output formats cannot support all these commands. As such, the corresponding plugins should support as large a subset as possible. The PDF output format is likely to support the largest set of commands and so it should be used to examine implementation details.

PAGE FORMATTING COMMANDS

- C100 enable pages to have unprinted areas around text, for bookbinding, trimming, and specific printing purposes
- C101 user-specified page size. (default: letter size)
- C101a Provide a large list of predefined sizes.
- C101b Allow page size to be specified on command line. (Permits printing of the same file in multiple page sizes without need to modify the Platypus file.)
- C102 page margins (default: 1 inch)
- C103 enable mirrored margins for left/right pages
- C104 user-specifiable page columns (default: 1, full width between margins)
- C105 ability to change columns on the fly
- C106 ability to specify column gutters
- C107 ability to mirror column layout on left/right pages

- C108 page color
- C108a provide 50 (or so) predefined page colors
- C109 new page
- C110 insert numerous blank pages
- C111 Headers for left and right pages
- C112 Headers must have three fields located at: outer margin, middle, inner margin
- C113 Any and all header fields can be left blank
- C114 Various format commands for supporting elements:
 - C114a Underline
 - C114b Overline
 - C114c Decorative character beside page number
- C115 Ability to disable printing header for any and all pages
- C116 Footers for left and right pages
- C117 Footers must have three fields located at: outer margin, middle, inner margin
- C118 Footers and all footer fields can be left blank
- C119 Various format commands for supporting footer elements, including:
 - C119a Underline
 - C119b Overline
 - C119c Decorative character beside page number
- C120 Ability to disable printing footer for any and all pages

PARAGRAPH FORMATTING COMMANDS

- C131 new paragraph
- C132 insertion of space at start of paragraph
- C133 skip line at beginning of paragraph
- C134 Alignment of paragraph to left margin
- C135 Alignment of paragraph to right margin
- C136 Alignment of paragraph to center
- C137 Justified alignment (left and right margins simultaneously)

- C138 Indentation of left side of paragraph
- C139 Indentation of right side of paragraph
- C140 Bulleted list, using user-specified bullet characters (default: round, centered dot)
- C141 Bulleted list, using numbers
- C141a Bulleted list, using numbers alternating between Arabic and Roman digits
- C142 Bulleted list, using letters
- C142a Bulleted list, using letters alternating between capitals and lower case
- C143 User-specified leading (spacing between lines) (default: 14pt)
- C143a Default leading based on size of text font
- C144 Special paragraphs for each of these sectional elements:
 - Book
 - Part
 - Chapter
 - Section
 - Sub-section
 - Paragraph
- C145 Automatic numbering of paragraphs
- C146 Automatic numbering of lines of text
- C146a Restart line numbering at 1 on new page
- C146b Continuous numbering of lines through new page
- C147 Footnotes (automatically numbered)
- C147a User-designated footnote numbering scheme: letters, numbers, symbols
- C148 End notes (at end of chapter)
- C149 End notes (at end of document)

NON-FONT TEXT FORMATTING COMMANDS

- C160 Hyphenation for any size block of text, including the entire document
- C161 Use of external language dictionaries for hyphenation

- C162 Specification of language hyphenation dictionaries via Platypus configuration file (to allow additional dictionaries to be added without rebuilding Platypus).
- C163 Kerning of letter pairs, larger blocks of letters, including a full paragraph.
- C164 Superscript (letter, number, whole words)
- C165 Two levels of superscripts
- C166 Subscript (letter, number, whole words)
- C167 Two levels of subscripts
- C168 Standard ligatures (on/off)
- C169 Non-standard ligatures
- C169a AE
- C169b OE

FONT FORMATTING COMMANDS

Note: Some of these command, strictly speaking are not part of the font or typeface specification, but they are included here for convenience.

- C181 Text size (default: 12pt)
- C182 Text weight (bold, not bold) (default: not bold)
- C183 Text color (default: black)
- C184 Italics (default: off)
- C185 Underline (default: off)
- C185a Underline with adjustable line thickness (default to be determined)
- C185b Underline with adjustable underline location (default to be determined)
- C186 Underline where whitespace is not underlined
- C187 Double underline
- C188 Overbar
- C189 Italics
- C190 Typeface selection
- C191 Support: OpenType, TrueType, Adobe AFM

- C192 Alternate fonts (note: primarily for use in HTML)
- C193 Ability to save and restore the current font (intended for brief font changes: save current font->change font->restore previous font)
- C194 Text background color
- C195 Strikethrough
- C195a User-definable strikethrough thickness (default to be determined, = to underline)
- C196a Strikethrough vertical location
- C196b Strikethrough color (default: = to text color)

OTHER DOCUMENT ELEMENT FORMATTING COMMANDS

- C210 Watermarks, based on images
 - C210a Watermarks above text
 - C210b Watermarks below text
 - C210c Partially transparent watermarks
- C211 Image (supported formats: GIF, JPG, PNG. Other formats to be decided)
 - C211a Location
 - C211b Scaling
 - C211c Ability to span multiple pages
 - C211d User-specifiable flow of text around images
- C212 Captions for images
 - C212a Automatic numbering for captions
- C220 Tables [NOTE: Requirements not yet formulated]
- C240 Table of Contents
 - C241 Specifying in-text which elements to include in Table of Contents
 - C241a Specify format of Table of Contents entries, including indentation
 - C242 Leader between Table entry and page number
- C243 Index
 - C244 Ability to specify levels of index
 - C245 Ability to specify items to be indexed

- C245a Ability to specify alternate text to appear in the index
- C245b Ability to specify special formatting for entry text in the index
- C246 User-definable formatting index sections (generally divided by starting letter)
- C247 List of figures
- C248 User definable leader in List of Figures between figure number, title, and page number
- C260 Forward/backward references to other text elements
- C260a Automatic removal of references if destination does not exist
- C261 URLs, with text hiding the hyperlink data (as commonly done in HTML)
- C261a User-definable formatting for making hyperlink stand out
- C270 Marginal text.
- C270a Marginal text formatting as any other text, save that it must fit in the margin space.
- C270b Tie location of marginal text to word(s) or element(s) in document
- C271 Tabs (unlimited number supported)
- C272 Decimal alignment of numerics
- C290 Graphics generation
- C291 Graphics lines
- C292 Graphics shapes
- C293 Irregular graphics shapes with path and boundaries specified
- C294 Shading of graphics
- C300 Formulas (Note: requirements still to be formulated)
- C361 Barcodes supporting
- C361a

ACTION COMMANDS

- C401 Define string macro
- C402 Define numeric macro (uses large decimal)
- C403 Define rounding for numeric macros (uses + .5 of rightmost digit, then truncation)

- C404 Dereference/print string macro
- C405 Dereference/print numeric macro
- C405a Specify format for printed numeric macro
- C406 Debugging commands
- C406a Dump settings for:
 - Document
 - Page
 - Paragraph
 - Text
 - All user macros
 - All system macros
- C407a Dump of setting includes:
 - Name of setting
 - Value of setting
 - Line in document where its value was last set. (Line 0 = default)
- C420 Script language [Exact requirements still to be determined.]
- C440 Code sections
- C440a Default formatting of code text
- C441 Line numbering of code listing (default: off)
- C441a Numbers for every 1, 5, or 10 lines; with or without hash marks in between (default: without)
- C441b Line numbering restarts a 1 for each listing or continues from previous line number.
- C442 Language-sensitive line breaks for:
 - Java
 - C/C++
 - Platypus
 - XML
- C443 Other programming languages can be added using plugins.

- C444 Line coloring for comments and strings in code listings
- C445 Limited interpretation of commands for code listings (to avoid interpreting valid code as Platypus commands)
- C445a Allow color specification for code in listings
- C445b Allow bold and italics in code listing
- C445c Allow for insertion of numbered bullets in listings (right side only)
- C446 Enable code sections to be read from external code files
- C470 Templates: Allow placeholders to be filled in by data read from file
- C471 Templates [Further requirements to be defined.]
- C491 User-defined horizontal space skip
- C492 User-defined vertical space skip

SPECIAL CHARACTERS

Note: Due to the vast number of possible special characters, they are listed here in rough groupings. Suffice it to say that eventually Platypus will have more special characters than any other typesetting or layout language. Further note: all Unicode characters will be supported by a Unicode escape command. The following special characters are provided for convenience.

- C600 Unicode numbering specification

LANGUAGES

All special characters for:

- C601 English
- C602 French (English + 34 characters)
- C603 German
- C604 Spanish (English + 16 characters)
- C605 Greek
- C606 Serbo-Croatian
- C607 Polish
- C608 Russian (Cyrillic) (47 characters)
- C609 Icelandic

- C610 Czech
- C615 Japanese Katakana (left to right) (96 characters)
- C616 Japanese Hiragana (left to right) (93 characters)
- C617 Hebrew (Note: for single words or short phrases) (27 characters)
- C620 International phonetic alphabet (approx. 75 glyphs; exact number varies)
- C621 International Morse code (36 standard, 15 extended characters)
- C631 Accents and diacritics with overstrike

SPECIAL PUNCTUATION

- C641 Spaces
 - C641a Non-breaking space
 - C641b Thin space
 - C641c en-space
 - C641d em-space
- C642 Dashes and hyphens
 - C642a hyphen
 - C642b minus sign
 - C643c en-dash
 - C643d em-dash
- C644 single quotation marks
 - C644a left single quotation mark
 - C644b right single quotation mark
- C645 double quotation marks
 - C645a left double quotation mark
 - C645b right double quotation mark
- C646 guillemts
 - C646a left guillemets
 - C646b right guillemets
- C647 pilcrow

- C648 footnote symbols
 - C648a dagger
 - C648b double dagger
 - C648c croix de Lorraine
 - C648d section symbol
- C649 bullets
 - C649a round bullet (small and large)
 - C649b circle
 - C649c square (filled in)

MATHEMATICAL AND SCIENTIFIC

- C651 All binary operation symbols defined in LaTeX (36 symbols)
- C652 All relation symbols defined in LaTeX (40 symbols)
- C653 All arrow symbols defined in LaTeX (33 symbols)
- C654 All miscellaneous mathematical symbols defined in LaTeX (33 symbols)
- C655 All variable-sized mathematical symbols defined in LaTeX (14 symbols)
- C656 Fractions
 - C656a Vulgar fractions: half (1 char)
 - C656b Vulgar fractions: quarters (2 chars)
 - C656c Vulgar fractions: eighths (4 characters)
 - C656d Vulgar fractions: thirds (2 characters)
- C657 %
- C658 per thousand
- C659 degree sign
- C660 units of measure
 - C660a angstrom
- C661 Zodiacal signs (12 symbols)
- C662 Planetary and Solar symbols (approx. 30 signs)

- C663 Roman Numerals (26 characters)
- C664 Bulleted numbers (Black background, 99 characters)
- C665 Bulleted numbers(White background, 99 characters)

MISCELLANEOUS CHARACTERS

- C701 Card suits (8 symbols)
- C702 Chess pieces including chess notation
- C703 Dingbats
- C704 Currency symbols (22 symbols)
- C705 Copyright symbol
- C706 Trademark symbol
- C707 Registered trademark symbol

OTHER REQUIREMENTS

- C901 Fully internationalizable implementation (by use of resource files for all literals).