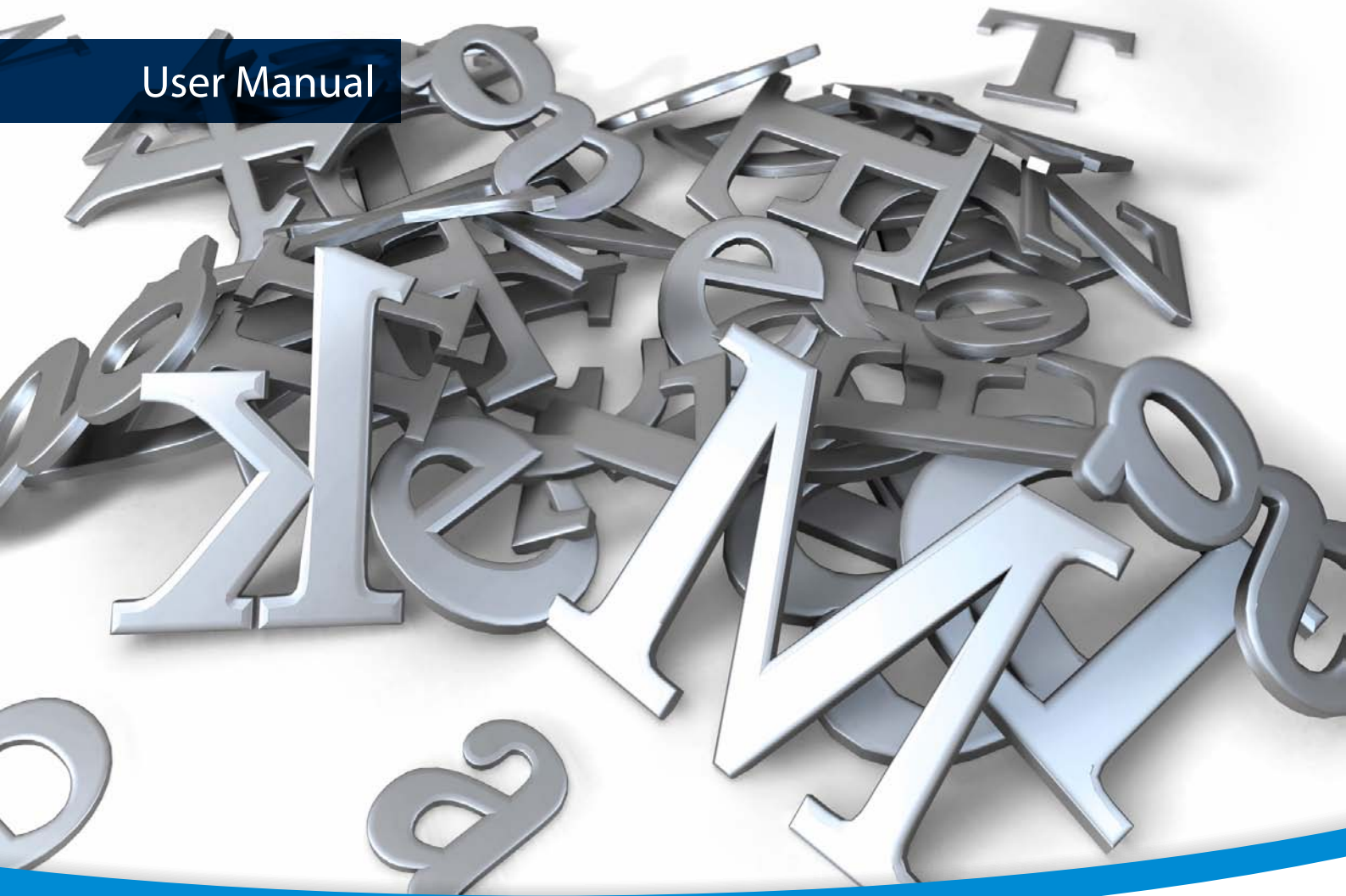


User Manual



4-Heights™ PDF Toolbox SDK

For Java

Version 2.4.0



Contents

1	Introduction	3
1.1	Description	3
1.2	Functions	3
1.2.1	Features	3
1.2.2	Formats	6
1.2.3	Conformance	7
1.3	Interfaces	7
1.4	Operating Systems	7
2	Installation	8
2.1	General Installation Steps	8
2.2	Java Specifics	8
2.3	Color Profiles	9
2.3.1	Default Color Profiles	9
2.3.2	Get Other Color Profiles	9
2.4	Fonts	10
2.4.1	Font Cache	10
2.5	Special Directories	10
2.5.1	Directory for temporary files	10
2.5.2	Cache Directory	10
2.5.3	Font Directories	11
2.6	License Keys	11
2.7	Un-Install, Install a New Version	11
3	User's Guide	12
3.1	General Concepts	12
3.1.1	Document model	12
3.1.2	Copying instead of modification	12
3.1.3	Differentiation between object creation and use	12
3.1.4	Generator Objects	13
3.1.5	Garbage collection and closing objects	13
3.2	Thread safety	13
3.2.1	Garbage Collection and Finalizer	14
3.3	The PDF Graphics Model	14
3.3.1	Coordinate System	14
3.3.2	Affine Transformations	14
3.4	Annotations and Form Fields	15
3.4.1	Form Fields	15
	Creating Form Fields	16
	Filling Form Fields	17
4	Java Interface	19
4.1	AutoCloseable Objects	19
4.2	Properties	19
4.3	Error handling	19
4.4	Streams	19
4.5	Lists	19
4.6	Enumerables	19
4.7	Maps	20

5	Version History	21
5.1	Version 2	21
6	Licensing, Copyright, and Contact	22

1 Introduction

1.1 Description

The 4-Heights™ PDF Toolbox SDK is a component to create, extract, assemble, and modify PDF documents.

This product is the successor of the 3-Heights™ PDF Toolbox API. For the task of migrating existing projects from \ThreeHeights to \FourHeights, PDF Tools offers an additional migration guide and a source code updating script. Please contact us for more information.

1.2 Functions

1.2.1 Features

Document assembly

- Copy pages from existing PDFs
- Copy annotations, form fields, links, logical structure, destinations, outlines, layers
- Flatten annotations, form fields, signatures
- Optimize resources
- Crop and rotate pages
- Free composition of content: overlays, underlays, stamps, transformations
- Encryption: user password, owner password, permissions
- Copy and modify document metadata
- Copy and modify page metadata
- Add embedded files and associated files
- Get and set OpenAction destination

Generation

Document Level

- Create pages
- Create form fields
 - General text fields and comb text fields
 - Check boxes
 - Radio button groups
 - List boxes
 - Combo boxes
- Create new outline items and insert them at any position in the tree
- Destinations: Named and direct destinations in the same document
- Configure viewer settings

Page Content Level

- Create new PDF content from scratch

- Apply content to existing pages

Colors

- Device colors: RGB, CMYK and grayscale
- ICC color profiles
- Transparency: alpha and blend mode

Paths

- Single and multi-segment lines
- Rectangle, circle, Bezier curves, ellipse, arc, pie
- Filling, stroking, clipping and combinations thereof
- Line width, cap, join, dash array, dash phase and miter limit
- Inside rule: nonzero winding rule, even/odd rule

Text

- Font size, character spacing, word spacing
- Horizontal scaling, leading, rise
- Enables simple text layouting
- Standard PDF fonts, installed fonts
- Font metrics: italic angle, ascent, descent, cap height, character width
- Unicode characters
- Text stroke: line width, line join and dashes
- Fill and stroke text, invisible text
- Use text as clipping path

Images

- Bi-level: CCITT G3, G3 2D and G4, Flate, LZW, Packbits, uncompressed
- 4 bit and 8 bit grayscale: Flate, LZW, Packbits, JPEG and JPEG-6 (8 bit only), uncompressed
- RGB: Flate, JPEG and JPEG-6, LZW, Packbits, uncompressed

Transformations

- Translation
- Scaling
- Skewing (horizontal, vertical)
- Rotation

Annotations

- Web link annotations
- File attachment annotations
- Free text annotation
- Sticky note annotation
- Text stamp annotation
- Custom stamp annotation
- Circle annotation
- Square annotation

- Line annotation
- Poly line annotation
- Polygon annotation
- Free drawing (“ink”) annotation

Modification

Page Content

- Selective deletion of content elements (without tagging and layers)
- Geometric transformation of content elements (without tagging and layers)

Annotations

- Web link annotations’ target URIs
- Markup annotations’ location, creation/modification date, subject, author, content

Form Fields

- Deletion of fields and modification of field values for
 - General text fields and comb text fields
 - Check boxes
 - Radio button groups
 - List boxes
 - Combo boxes

Extraction

Document and Page

- Document information entries: title, author, subject, keywords, creator, producer, creation date, modification date
- Document XMP metadata
- Document encryption settings
- Embedded files
- Page bounding boxes: media box, crop box, bleed box, trim box, art box
- Page XMP metadata
- Outline item tree: Tree structure, item title, expanded/collapsed
- Destinations: Named and direct destinations in the same document
- Viewer settings

Content

Page’s and group’s content elements including

- Bounding box
- Affine transformation

As either of the following:

- Group element
- Image element

- Width and height in pixel
- Bits per component
- Color space
- Image mask element
 - Width and height in pixel
 - Paint for filling the mask
- Path element
 - Alignment box
 - Fill parameters including paint and fill rule
 - Stroke parameters including line paint and line style
- Shading element
- Text element
 - Text fragments
 - Bounding box
 - Affine transformation
 - Unicode string
 - Fill parameters including paint and fill rule
 - Stroke parameters including line paint and line style

Annotations

- Annotations: location
- Mark-up annotation: type, location, creation/modification date, subject, author, content
- Custom stamp annotations: appearance
- Link annotations: location, target destination or URI
- Signature fields: name, location, reason, contact info, date, visibility

AcroForm Form Fields

- Form field identifiers, export names and user names, including form field hierarchy
- Form field export and display content of:
 - Push buttons
 - Check boxes
 - Radio button groups
 - General text fields and comb text fields
 - List boxes
 - Combo boxes

1.2.2 Formats

Supported PDF Formats

- PDF 1.x (PDF 1.0, . . . , PDF 1.7)
- PDF 2.0
- PDF/A-1, PDF/A-2, PDF/A-3

Supported Image Formats

- BMP
- DIB

- JPEG
- JPEG2000
- JBIG2
- PNG
- GIF
- TIFF

Supported Font Formats

- Type1
- TrueType
- OpenType
- OpenType (CFF)

1.2.3 Conformance

Standards:

- ISO 32000-1 (PDF 1.7)
- ISO 32000-2 (PDF 2.0)
- ISO 19005-1 (PDF/A-1)
- ISO 19005-2 (PDF/A-2)
- ISO 19005-3 (PDF/A-3)

1.3 Interfaces

The following interfaces are available:

- C
- Java
- .NET Framework
- .NET Core¹

1.4 Operating Systems

The 4-Heights™ PDF Toolbox SDK is available for the following operating systems:

- Windows Client 7+ | x86 and x64
- Windows Server 2008, 2008 R2, 2012, 2012 R2, 2016, 2019 | x86 and x64
- Linux:
 - Red Hat, CentOS, Oracle Linux 8+ | x64
 - Fedora 29+ | x64
 - Debian 10+ | x64
 - Other: Linux kernel 2.6+, GCC toolset 4.8+, glibc 2.27+ | x64
- macOS 10.10+ | x64

‘+’ indicates the minimum supported version.

¹ Limited supported OS versions. [Operating Systems](#)

2 Installation

2.1 General Installation Steps

The “4-Heights™ PDF Toolbox SDK For Java” comes as a ZIP archive. The installation of the software requires the following steps:

1. Ensure that your system matches one of the supported [Operating Systems](#).
2. Log in to your account on <https://www.pdf-tools.com/> and download the ZIP archive and the license key for “PDF Toolbox SDK”. (See also [License Keys](#).)
3. Unzip the archive to a local directory, e.g, on Windows: C:\Program Files\PDF Tools AG\, on Linux: /opt/pdftools.com/. This creates the following subdirectories:

Subdirectory	Description
bin	Contains the runtime executable binaries for all supported platforms: <ul style="list-style-type: none">▪ linux/libPdfToolbox.so for 64 bit Linux▪ macOS/libPdfToolbox.dylib for 64 bit macOS▪ Win32\PdfToolbox.dll for 32 bit Windows▪ x64\PdfToolbox.dll for 64 bit Windows
doc	Contains documentation
jar	Contains Java archive file com.pdf_tools.fourheights.pdftoolbox.jar.

4. On Windows, you may want to add the bin\x64 or bin\Win32 sub-directory to the %PATH% environment variable.

On Linux, you may want to create a link to the shared library from one of the standard library directories, e.g:

```
ln -s /opt/pdf-tools.com/bin/linux/libPdfToolbox.so /usr/lib
```

(See [Java Specifics](#) for an alternative.)

5. On macOS, the library must have the extension .jnilib for use with Java. You may want to create a file link for this purpose, e.g:

```
ln libPdfToolbox.dylib libPdfToolbox.jnilib
```

6. Ensure that your platform meets the requirements regarding color profiles. (See [Color Profiles](#).)
7. Ensure that your platform meets the requirements regarding fonts. (See [Fonts](#).)
8. Ensure that the cache directory exists and is writable. (See [Special Directories](#).)

2.2 Java Specifics

The 4-Heights™ PDF Toolbox SDK requires Java version 7 or higher.

For compilation and execution When using the Java interface, the Java archive jar\com.pdf_tools.fourheights.pdftoolbox.jar needs to be on the CLASSPATH. This can be done by either adding it to the environment variable CLASSPATH, or by specifying it using the switch `-classpath` or `-cp`:

```
javac -cp ".;C:\Program Files\PDF Tools AG\jar\com.pdf_tools.fourheights.pdftoolbox.jar" ^
sampleApplication.java
```

For execution Additionally the library (on Windows: PdfToolboxAPI.dll) needs be in one of the system's library directories² or added to the Java system property `java.library.path`. This can be achieved by either adding this system property dynamically at program startup before using the API, or by specifying it using the switch `-Djava.library.path` when starting the Java VM. Choose the correct subdirectory (x64 or Win32 on Windows) depending on the platform of the Java VM³.

```
java -cp ".;C:\Program Files\PDF Tools AG\com.pdf_tools.fourheights.pdftoolbox.jar" ^
-Djava.library.path=C:\Program Files\PDF Tools AG\bin\x64" sampleApplication
```

Note that on Linux or macOS, the path separator usually is a colon and hence the above changes to something like:

```
... -cp ".:path/to/com.pdf_tools.fourheights.pdftoolbox.jar" ...
```

2.3 Color Profiles

If no color profiles are available, default profiles for both RGB and CMYK are generated on the fly by the 4-Heights™ PDF Toolbox SDK.

2.3.1 Default Color Profiles

If no particular color profiles are set default profiles are used. For device RGB colors a color profile named "sRGB Color Space Profile.icm" and for device CMYK a profile named "USWebCoatedSWOP.icc" are searched for in the following directories:

Windows

1. %SystemRoot%\System32\spool\drivers\color
2. directory `Icc`, which must be a direct sub-directory of where the PdfToolboxAPI.dll resides.

Linux and macOS

1. `$PDF_ICC_PATH` if the environment variable is defined
2. the current working directory

2.3.2 Get Other Color Profiles

Most systems have pre-installed color profiles available, for example on Windows at %SystemRoot%\system32\spool\drivers\color\. Color profiles can also be downloaded from the links provided in the directory `bin\Icc\` or from the following websites:

- <http://www.pdf-tools.com/public/downloads/resources/colorprofiles.zip>
- <http://www.color.org/srgbprofiles.html>
- https://www.adobe.com/support/downloads/iccprofiles/iccprofiles_win.html

² On Windows defined by the environment variable `PATH` and e.g. on Linux defined by `LD_LIBRARY_PATH`.

³ If the wrong data model is used, there is an error message similar to this: "Can't load IA 32-bit .dll on a AMD 64-bit platform"

2.4 Fonts

When text is created by the 4-Heights™ PDF Toolbox SDK, all fonts from the [Font Directories](#) can be used.

Note that on Windows when a font is installed it is by default installed only for a particular user. It is important to either install fonts for all users, or make sure the 4-Heights™ PDF Toolbox SDK is run under that user and the user profile is loaded.

2.4.1 Font Cache

A cache of all fonts in all [Font Directories](#) is created. If fonts are added or removed from the font directories, the cache is updated automatically.

In order to achieve optimal performance, make sure that the cache directory is writable for the 4-Heights™ PDF Toolbox SDK. Otherwise the font cache cannot be updated and the font directories have to be scanned on each program startup.

The font cache is created in the subdirectory <CacheDirectory>/Installed Fonts of the [Cache Directory](#).

2.5 Special Directories

2.5.1 Directory for temporary files

This directory for temporary files is used for data specific to one instance of a program. The data is not shared between different invocations and deleted after termination of the program.

The directory is determined as follows. The product checks for the existence of environment variables in the following order and uses the first path found:

Windows

1. The path specified by the %TMP% environment variable.
2. The path specified by the %TEMP% environment variable.
3. The path specified by the %USERPROFILE% environment variable.
4. The Windows directory.

Linux and macOS

1. The path specified by the \$PDFTMPDIR environment variable.
2. The path specified by the \$TMP environment variable.
3. The /tmp directory.

2.5.2 Cache Directory

The cache directory is used for data that is persisted and shared between different invocations of a program. The actual caches are created in subdirectories. The content of this directory can safely be deleted to clean all caches.

This directory should be writable by the application, otherwise caches cannot be created or updated and performance will degrade significantly.

Windows

- If the user has a profile:
%LOCAL_APPDATA%\PDF Tools AG\Caches

- If the user has no profile:
<TempDirectory>\PDF Tools AG\Caches

Linux and macOS

- If the user has a home directory:
~/ .pdf-tools/Caches
- If the user has no home directory:
<TempDirectory>/pdf-tools/Caches

where <TempDirectory> refers to the [Directory for temporary files](#).

2.5.3 Font Directories

The location of the font directories depends on the operating system. Font directories are traversed recursively in the order as specified below.

If two fonts with the same name are found, the latter one takes precedence, i.e. user fonts will always take precedence over system fonts.

Windows

1. %SystemRoot%\Fonts
2. User fonts listed in the registry key \HKEY_CURRENT_USER\Software\Microsoft\Windows NT\CurrentVersion\Fonts. This includes user specific fonts from C:\Users\<user>\AppData\Local\Microsoft\Windows\Fonts and app specific fonts from C:\Program Files\WindowsApps
3. directory Fonts, which must be a direct sub-directory of where PdfToolboxAPI.dll resides.

macOS

1. /System/Library/Fonts
2. /Library/Fonts

Linux

1. /usr/share/fonts
2. /usr/local/share/fonts
3. ~/.fonts
4. \$PDFFONTDIR or /usr/lib/X11/fonts/Type1

2.6 License Keys

The 4-Heights™ PDF Toolbox SDK can only be used with a valid license key. This key must be set programmatically by using the `Sdk.Initialize` method prior to making any calls to the library. You can download your license key from your account on <https://www.pdf-tools.com/>.

For licensing questions, please contact pdfsales@pdf-tools.com.

2.7 Un-Install, Install a New Version

Un-installation is done by undoing the steps in [General Installation Steps](#), specifically:

- Remove un-zipped files
- Revert environment variables if necessary
- Remove links if necessary

An update to a new version is done by a drop-in replacement of the existing files. Make sure that all files are updated consistently.

3 User's Guide

3.1 General Concepts

3.1.1 Document model

The document model of the PDF Toolbox SDK consists of two different types of objects:

Structure objects define the structure of the document, such as `Document`, `Page` or `Content`.

Graphics resources can be used to draw content with a `ContentGenerator`. Examples are `Image`, `Font` or `ColorSpace`.

All objects in the document model are bound to a specific document. They can only be used in the context of the document for which they were created using their `Create` or `Copy` method.

The objects of the document model are all stateless. Where a stateful interface is useful, it is provided by an external [generator](#) or extractor, which is not considered part of the document model.

3.1.2 Copying instead of modification

The PDF Toolbox SDK does not allow in-place modification of documents. Instead the content is copied into a new document, while performing the necessary changes.

To copy objects from a source document into a target document, the object's static `Copy` method is called with the **target** document as first argument.

This concept allows the processing of very large files without consuming much memory: The content of the input document is only read on demand and any modifications can be directly stored in the output file.

3.1.3 Differentiation between object creation and use

To provide a uniform interface, many operations are divided into two steps:

1. Create (or copy) the object
2. Use the object

This separation allows to provide multiple variants for both steps, without having a "combinatorial explosion" of methods.

Step 1: Create

In the first step, the object is created in the target document or copied from the source document to the target document.

After creating, the object is associated with the document, but not yet used. This means, that copying or creating an object may change the size of the target file, but logically, the PDF is still unchanged.

Examples are the following methods:

- `Page.Create`
- `Font.Create`
- `Page.Copy`
- `PageList.Copy`

- `ColorSpace.Copy`
- `Metadata.Copy`
- `ContentElement.Copy`

Step 2: Use

The associated object can then be used in the target document.

This second step is often more lightweight than the first step, since all the necessary copying is already done.

Examples are the following methods of a `ContentGenerator`:

- `PaintImage`
- `PaintGroup`
- `AppendContentElement`

or the `PageList.Add` method.

3.1.4 Generator Objects

Some objects in a PDF consist of a list or stream of operations that operate on an internal state:

- Content streams
- Text objects
- Path objects

Since all data objects in the PDF Toolbox SDK are stateless, a (simplified) stateful interface is provided by so called generator interfaces:

- Content objects can be modified with a `ContentGenerator`.
- Path objects can be modified with a `PathGenerator`.
- Text objects can be modified with a `TextGenerator`.

Generator objects must always be closed explicitly, before the generated object can be used.

3.1.5 Garbage collection and closing objects

Every interface object is considered being a resource that needs to be closed after use. Most objects are closed automatically, at the latest when the owning document is closed, in C# and Java possibly earlier by the garbage collector.

In addition to `Document` objects, the “generator” objects `ContentGenerator`, `PathGenerator`, and `TextGenerator` must be closed, lest the generated objects are incomplete.

3.2 Thread safety

The PDF Toolbox SDK is generally thread-safe with one exception:

A document may only be accessed in one thread concurrently, including all sub-objects.

Note that almost all objects are directly or indirectly associated with a document.

Note: Methods that copy from a source to a target document have to access both documents. The thread safety rules not only apply to the target document, but also to the source document. This means that concurrently copying from the same source document is not allowed.

3.2.1 Garbage Collection and Finalizer

Object finalization is thread-safe with one exception:

The finalizer of the Document is not thread-safe regarding access to its sub-objects.

Sub-objects do not retain their associated document object. If all references to an open document go out of scope, the document finalizer will eventually be running and the document will be closed.

Explicitly accessing (even closing) any sub-object while the document finalizer is running is **not safe!**

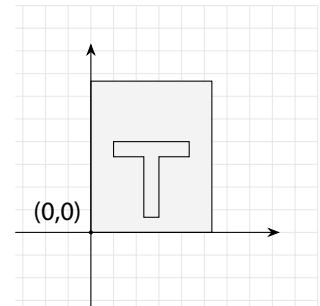
3.3 The PDF Graphics Model

3.3.1 Coordinate System

PDF coordinates are measured from bottom to top, in contrast to many other coordinate systems used in informatics.

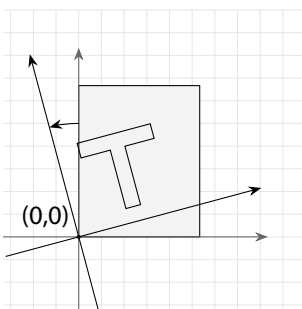
For the sake of simplicity, all coordinates used in the PDFToolbox SDK are normalized, such that the point (0,0) denotes the lower left corner of the visible page (crop box).

The internal **Rotate** attribute of a PDF page is not exposed at the API. Instead, all coordinates are assumed to refer to the already rotated page.

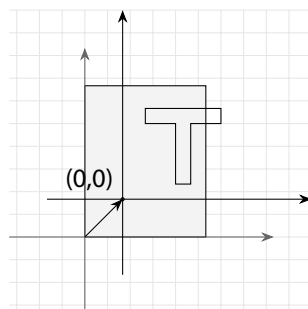


3.3.2 Affine Transformations

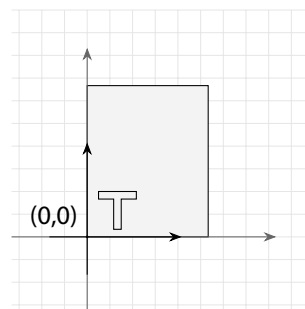
Affine transformations can be used to rotate, move, scale, or otherwise skew any page content.



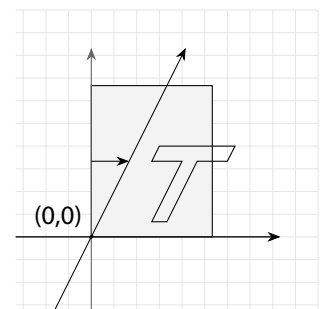
Rotate



Move



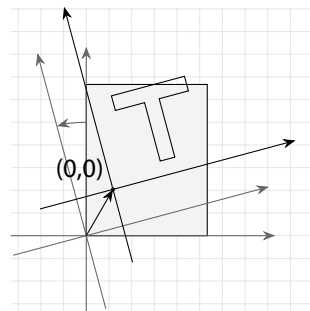
Scale



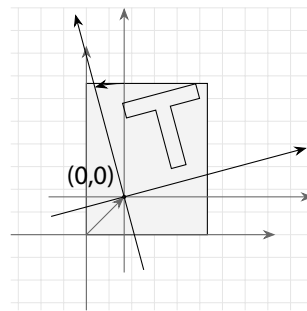
Skew

Transformations always affect the coordinate system as a whole. All **following** graphics operations are executed in the transformed coordinate system, including additional transformations.

This means, that the ordering how transformations are applied is important.

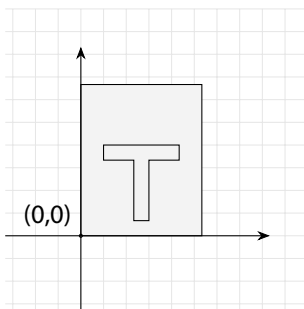


Rotate, then Move

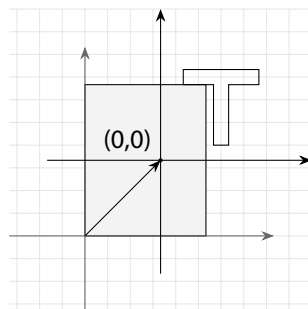


Move, then Rotate

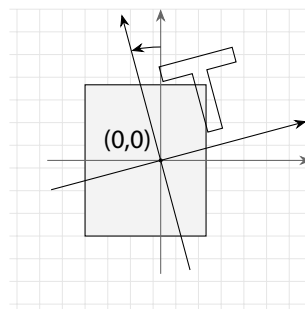
Example: Rotate around a certain point



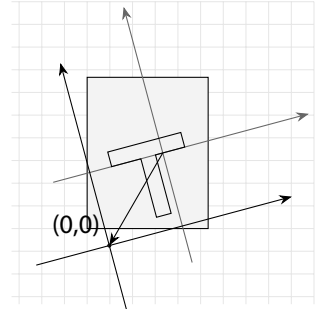
Start ..



.. Move to center point ..



.. Rotate ..



.. Move back to origin

3.4 Annotations and Form Fields

Annotations are elements that are not part of a page's content but are applied on top of a page. In contrast to ordinary page content, many annotation types are meant to behave interactively in a PDF viewer.

3.4.1 Form Fields

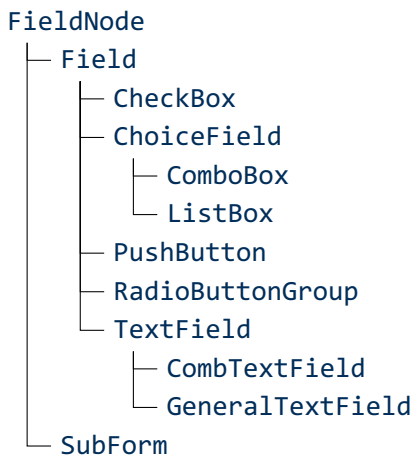
AcroForm form fields in a PDF consist of data structures that represent variable values, potentially to be modified by a user in a PDF viewing application.

Form fields are structured in a tree topology in which the `Document` acts as the tree root. The immediate child form field nodes thereof can be accessed via the `Document.FormFields` property.

Each child form field node in the tree can itself have more children. We name such a form field a "sub form". A child form field node that has no children is simply a "form field", of which the following concrete sub-types exist:

- General text field
- Comb text field
- Check box field
- Push button field
- Radio button field
- Combo box field
- List box field

In the PDF Toolbox SDK, the above types are modeled by classes that inherit from a base class `FieldNode` in the following way:



`FieldNodes` are contained in a `FieldNodeMap` in the form of key-value pairs for which the keys act as form field node **identifiers**. Such an identifier is not allowed to contain any full stops (“.”).

The **fully qualified identifier** of a form field node is defined as the concatenation of all its ancestor `SubForm`’s identifiers and its own identifier, separated by full stops (“.”), e.g. `MySubForm.MyField`. The fully qualified identifier of each form field node is unique within a document.

While the form field tree models the form’s data, the visual manifestations of form fields are managed by `Widgets` contained in `Page.Widgets`, of which each form field has at least one.

Creating Form Fields

In an output document (a document created with `Document.Create`), form fields can be created from scratch by means of the following methods:

- `CheckBox.Create`
- `ComboBox.Create`
- `CombTextField.Create`
- `GeneralTextField.Create`
- `ListBox.Create`
- `RadioButtonGroup.Create`
- `SubForm.Create`

A `PushButton` cannot be created.

After creating a `ChoiceField`, i.e. a `ComboBox` or a `ListBox`, `ChoiceItems` should be created and added with the `ChoiceField.AddNewItem` method.

After creating a `RadioButtonGroup`, new `RadioButtons` should be created and added with the `RadioButtonGroup.AddNewButton` method.

Each created form field must be added either to the document’s `Document.FormFields`, or to the `SubForm.Children` of a `SubForm`.

It is advisable to set all form field properties prior to creating any `Widgets`. Specifically, changing form field properties that affect the form field’s visual appearance fails with an `UnsupportedOperation` error if the form field has widgets.

For each form field, at least one `Widget` should be created using the method `RadioButton.AddNewWidget` for radio button groups, or the form field’s method `Field.AddNewWidget` for all other field types.

Finally, each created widget must be added to one of the `Page.Widgets` of any of the document’s `Pages`.

A page can either be created from scratch with the `Page.Create` method or it can be copied with `Page.Copy` from an input document (a document created with `Document.Open`). In the latter case, `PageCopyOptions.FormFields` must not be set to `FormFieldCopyStrategy.Copy` and `PageCopyOptions.UnsignedSignatures` not to `CopyStrategy.Copy`.

The combination of creating form fields and copying form fields or unsigned signatures via `Page.Copy` or `PageList.Copy` (with `PageCopyOption` argument in which either the `PageCopyOptions.FormFields` property is set to `FormFieldCopyStrategy.Copy` or the `PageCopyOptions.UnsignedSignatures` property is set to `CopyStrategy.Copy`) is not supported. Specifically:

- Once `Page.Copy` or `PageList.Copy` has been called with `PageCopyOptions.FormFields` set to `FormFieldCopyStrategy.Copy` or `PageCopyOptions.UnsignedSignatures` set to `CopyStrategy.Copy`, any subsequent call to any of the form field creation methods fails with an `IllegalState` error.
- Once any of the form field creation methods has been called, any subsequent call to `Page.Copy` or `PageList.Copy` with `PageCopyOptions.FormFields` set to `FormFieldCopyStrategy.Copy` or `PageCopyOptions.UnsignedSignatures` set to `CopyStrategy.Copy` fails with an `IllegalArgument` error.

Filling Form Fields

Filling a form means that the values (field content) of form fields are modified. Depending on the field type this implies the following:

- `TextField`: modify text using `TextField.Text`.
- `CheckBox`: modify the state using `CheckBox.Checked`.
- `RadioButtonGroup`: modify the choice using `RadioButtonGroup.ChosenButton`.
- `ComboBox`: modify choice using `ComboBox.ChosenItem` or `ComboBox.EditableItemName`.
- `ListBox`: modify choice using `ListBox.ChosenItems`.

In order to use the PDF Toolbox SDK for filling out the values of form fields in a PDF, the following procedure must be followed:

1. An input document is opened with `Document.Open` and an output document is created with `Document.Create`.
2. Before copying pages, the form fields must be copied from the input document to the output document as follows:
 - a. Access the form field node map of the input and the output document via `Document.FormFields`.
 - b. Copy each form field node found in the input to the output document using the `FieldNode.Copy` method. Copying `SubForms` automatically copies their children. (Note that the copied form fields have no widgets yet.)
 - c. The value of a copied form field can be modified here or later in [Step 3](#).
 - d. Add each copied form field node to the output document's field node map `Document.FormFields`, preferably using the same key as used in the input document's field node map.
3. The output document's form field nodes can be accessed, e.g. using `FieldNodeMap.Lookup`, to modify form field values.
4. Copy all pages with the `PageList.Copy` method. Hereby, the `PageCopyOptions.FormFields` property in the `PageCopyOptions` argument must be set to `FormFieldCopyStrategy.CopyAndUpdateWidgets` and the `PageCopyOptions.UnsignedSignatures` must be set to `CopyStrategy.Remove` or `CopyStrategy.Flatten`. In this step, the `Widgets` of input form fields are copied to the output form fields and automatically updated to reflect the new form field values. (As soon as a form field has `Widgets`, its value cannot be modified anymore.)

The combination of filling form fields and copying form fields via `Page.Copy` or `PageList.Copy` (with `PageCopyOptions` argument in which `CopyOptions.FormFields` is set to `FormFieldCopyStrategy.Copy`) is not supported. Specifically:

- Once `Page.Copy` or `PageList.Copy` has been called with `PageCopyOptions.FormFields` set to `FormFieldCopyStrategy.Copy`, any subsequent call to `FieldNode.Copy` fails with an `IllegalState` error.
- Once `FieldNode.Copy` has been called, any subsequent call to `Page.Copy` or `PageList.Copy` with `PageCopyOptions.FormFields` set to `FormFieldCopyStrategy.Copy` or `PageCopyOptions.UnsignedSignatures` set to `CopyStrategy.Copy` fails with an `IllegalArgument` error.

4 Java Interface

4.1 AutoCloseable Objects

Objects that must be closed explicitly implement the `AutoCloseable` interface. Instead of calling `close()` directly, it is recommended to use the “try-with-resources” statement:

```
try (Document document = ...) {  
    ...  
} // document.close() is called implicitly here
```

See also [Garbage collection and closing objects](#).

4.2 Properties

Properties are modeled with setter and getter methods.

4.3 Error handling

Errors are reported using exceptions.

The following logic errors are mapped to the corresponding native runtime exception classes and are not checked:

IllegalArgumentException maps to `java.lang.IllegalArgumentException`

IllegalState maps to `java.lang.IllegalStateException`

UnsupportedOperation maps to `java.lang.UnsupportedOperationException`

Additionally, the following infrastructure error is mapped:

IO maps to `java.io.IOException`

The remaining errors are modeled using exception classes that inherit from a base class `PdfToolboxException`.

4.4 Streams

The native stream interfaces cannot be used, because they are lacking two important features:

- The PDF file format is based on random access. Native Java streams have only limited support for this.
- The ability to read from an output stream is crucial for processing large files (See [Chapter 3.1.2](#)).

Instead we provide a custom stream interface `com.pdf_tools.fourheights.pdftoolbox.Stream`, which has a similar interface as `java.io.RandomAccessFile`.

An implementation for files is provided, backed by `java.io.RandomAccessFile`.

4.5 Lists

Lists implement the native Java list interface `java.util.List`.

4.6 Enumerables

Enumerables (lists that only allow iterating) implement the native Java iterator interface `java.util.Iterable`.

4.7 Maps

Maps implement the native Java map interface `java.util.Map`.

5 Version History

5.1 Version 2

Changes in Version 2

- **Changed** support for sub-lists of page lists: The subList method now returns an object of type Pdf.PageList.
- **New** support for copying page lists.

6 Licensing, Copyright, and Contact

PDF Tools AG is a world leader in PDF (Portable Document Format) software, delivering reliable PDF products to international customers in all market segments.

PDF Tools AG provides server-based software products designed specifically for developers, integrators, consultants, customizing specialists and IT-departments. Thousands of companies worldwide use our products directly and hundreds of thousands of users benefit from the technology indirectly via a global network of OEM partners. The tools can be easily embedded into application programs and are available for a multitude of operating system platforms.

Licensing and Copyright The 4-Heights™ PDF Toolbox SDK is copyrighted. This user's manual is also copyright protected; It may be copied and given away provided that it remains unchanged including the copyright notice.

Contact

PDF Tools AG
Brown-Boveri-Strasse 5
8050 Zürich
Switzerland
<http://www.pdf-tools.com>
pdfsales@pdf-tools.com