

Optical Character Recognition Program
ABBYY® FineReader®

Version 8.0
User's Guide

Information in this document is subject to change without notice and does not bear any commitment on the part of ABBYY. The software described in this document is supplied under a license agreement. The software may only be used or copied in strict accordance with the terms of the agreement. It is a breach of the "On legal protection of software and databases" law of the Russian Federation and of international law to copy the software onto any medium unless specifically allowed in the license agreement or nondisclosure agreements.

No part of this document may be reproduced or transmitted in any form or by any means, electronic or other, for any purpose, without the express written permission of ABBYY.

© 2005 ABBYY Software. All rights reserved.

© 1987–2003 Adobe Systems Incorporated. Adobe® PDF Library is licensed from Adobe Systems Incorporated.

Microsoft Reader Content Software Development Kit © 2004 Microsoft Corporation, One Microsoft Way, Redmond, Washington 98052–6399 U.S.A. All rights reserved.

Fonts Newton, Pragmatica, Courier © 2001 ParaType, Inc. Font OCR–v–GOST © 2003 ParaType, Inc.

© 1999–2000 Image Power, Inc. and the University of British Columbia, Canada. © 2001–2002 Michael David Adams. All rights reserved.

ABBYY, the ABBYY Logo, Scan&Read, ABBYY FineReader are either registered trademarks or trademarks of ABBYY Software Ltd.

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

Microsoft, Outlook, Excel, PowerPoint, Windows are either registered trademarks or trademarks of Microsoft Corporation in the United States and/or other countries.

All other trademarks are the property of their respective owners.

Contents

Welcome!	4
What's New in ABBYY FineReader 8.0	5
Chapter 1 Working with ABBYY FineReader	7
Installing and Starting ABBYY FineReader.....	8
Acquiring the Image.....	9
Page Layout Analysis.....	16
Recognition.....	21
Checking and Editing Text.....	28
Saving into External Applications and Formats.....	33
Working with Batches.....	42
Automated Tasks.....	44
Chapter 2 ABBYY Screenshot Reader	51
Installing and Starting ABBYY Screenshot Reader.....	52
ABBYY Screenshot Reader Toolbar.....	52
Capturing Text and Tables from the Computer Screen.....	52
Making Screenshots.....	53
Additional Options.....	54
Chapter 3 ABBYY Hot Folder & Scheduling	55
Installing and Running ABBYY Hot Folder & Scheduling.....	56
ABBYY Hot Folder & Scheduling Main Window.....	56
Setting Up a Hot Folder.....	57
Hot Folder Log File.....	58
Additional Options for ABBYY Hot Folder & Scheduling.....	59
Appendix	60
Supported Document Saving Formats.....	61
Supported Image Formats.....	61
Hot Keys.....	62
Glossary.....	63

Welcome!

Thank you for purchasing ABBYY FineReader!

Electronic documents are becoming increasingly prevalent. However, business contracts, books and periodicals are still printed and millions of people use ABBYY FineReader to convert hard-copy documents into electronic formats.

ABBYY FineReader gives you the edge by providing full control over printed information: you can quickly transform any printed text or PDF file into an editable format and re-use their content.

ABBYY FineReader will help you:

- collect information from various sources and draw up a report
- edit a paper document or fax
- write an article, a thesis or a paper for publication
- publish newspaper and book clippings on the Web
- extract text from a PDF file and make changes to it

ABBYY FineReader is very easy to use. Even if you are a novice to optical character recognition, you will get results in a matter of minutes. And if you are an OCR professional, you can have full control over all the OCR settings and parameters. This User's Guide will introduce you to the features and commands of ABBYY FineReader and help you teach your computer to "read."

Welcome to OCR!

What's New in ABBYY FineReader 8.0

Compared to the previous version, ABBYY FineReader 8.0 introduces a variety of improvements and new features to increase your productivity when working with scanned documents, images, PDF files, and faxes.

This version features intelligent technology improvements, allowing for improved reading of images taken by digital cameras, more accurate recognition of low-resolution faxes and paper documents, better handling of document layouts, and enhanced security features in PDF files. New features like scheduled operation, recognition of screenshots and automation manager for processing recurring document tasks have been added to increase your productivity even more.

Detailed information on major product improvements and new features is given below. Features available only in ABBYY FineReader 8.0 Corporate Edition are marked accordingly.

Up to 30 percent accuracy improvement on low resolution documents and faxes

While ABBYY FineReader has traditionally delivered highly accurate recognition results on documents of good quality, there is still a possibility of documents of lower-than-expected quality arriving at your desk. Most often these are faxes or paper documents that were scanned at resolution lower than recommended for OCR. ABBYY FineReader 8.0 handles such documents better, delivering up to a 30 percent improvement in recognition accuracy.

Processing images taken by digital cameras

When you are on the go and no scanner is available, you may still capture documents with a digital camera and recognize them later on your desktop PC. Now ABBYY FineReader 8.0 includes new adaptive recognition technology for better OCR of camera images.

Security options for PDF files

The new version of ABBYY FineReader supports PDF security settings and allows you to set document Open and Permissions passwords as well as selecting other security options for PDF files. You can select RC4-based 40-bit or 128-bit encryption or the newest AES (Advanced Encryption Standard)-based 128-bit encryption.

Creating tagged PDF files

A new option for saving tagged PDF files in ABBYY FineReader 8.0 makes it possible to create PDF files that will be easier to read when used on devices with limited screen sizes, for example handheld devices.

Automation Manager

This new feature allows faster processing of repeated document tasks by grouping them into sets of consecutive operations that can then be called by one click of a button. Several predefined automated tasks are available and it is also possible to create your own customized automated tasks and share them with colleagues.

Support for hyperlinks

The new version recognizes hyperlinks, such as links to Web sites and e-mail addresses, and reconstructs them in output documents. You can also add new hyperlinks into recognized documents.

Fast mode recognition

With ABBYY FineReader 8.0 you can recognize documents 2–2 1/2 times faster using a new fast recognition mode. This mode is recommended for documents with simple layouts and good printing and scanning quality. For more complex documents, the accuracy mode should be preferred. However, recognition accuracy obtained in fast mode will be sufficient in many cases, for example when converting paper documents into searchable PDF files.

Saving to Microsoft Reader e-book (LIT) format

Now you can save recognition results in Microsoft Reader's LIT e-book format that makes such documents suitable for reading on handheld devices and PDAs.

Defining document-related properties

ABBYY FineReader 8.0 allows for defining additional document properties like Title, Author, Subject and Keywords, and saving this data to PDF, DOC/RTF, XLS, HTML, Word XML and LIT file formats. These properties can be used by the operating system and other software for indexing and search purposes.

Extended language and dictionary support

The total number of supported languages is now 179. Dictionary support and spell-check functions are available for 36 languages. Legal and medical dictionaries for the English and German languages are now included in the main English and German recognition dictionaries – there is no need to select specialized recognition languages to work with specialized text.

Opening multi-page PDF and TIFF files

If you do not need the entire document converted, you can open only selected pages of your multi-page PDF or TIFF files in ABBYY FineReader 8.0.

ABBYY Screenshot Reader

(available in ABBYY FineReader 8.0 Professional Edition after registration, available by default in ABBYY FineReader 8.0 Corporate Edition)

This simple and easy-to-use utility allows you to grab a part of the screen and recognize the text in the captured image. The utility also allows you to save captured screen areas to a file or to the clipboard.

ABBYY Hot Folder & Scheduling

(available only in ABBYY FineReader 8.0 Corporate Edition)

The previously available feature for automatically scanning folders for incoming images and processing these images has been extended in the new version of ABBYY FineReader. Scheduled processing capabilities have been added to allow you to utilize your computer for document conversion purposes when it is not occupied by your normal activities, for example at night time.

Chapter 1

Working with ABBYY FineReader

Chapter Contents:

- Installing and Starting ABBYY FineReader
- Acquiring the Image
- Page Layout Analysis
- Recognition
- Checking and Editing Text
- Saving into External Applications and Formats
- Working with Batches
- Automated Tasks
- Network Document Processing

Installing and Starting ABBYY FineReader

This chapter provides detailed instructions on installing ABBYY FineReader, outlines the system requirements of the program and offers instructions for installing the program on workstations and networks. ABBYY FineReader 8.0 includes a specialized installation program that automates the setup process. To ensure proper installation, always use the ABBYY FineReader CD-ROM for installation.

Software and Hardware Requirements

ABBYY FineReader 8.0 requires the following:

1. A PC with Intel® Pentium®/Celeron®/Xeon™, AMD K6/Athlon™/Duron™/Sempron™ or compatible processor (500 MHz or higher).
2. Operating System: Microsoft® Windows® Server 2003, Microsoft® Windows® XP, Microsoft® Windows® 2000. (To work with localized interfaces, corresponding language support is required.)
3. Memory: 128 MB RAM. In a multiprocessor system, an additional 16MB of RAM is required for each additional processor.
4. Hard disk space: 250MB for typical program installation and 100MB for program operation.
5. A TWAIN-compatible scanner, digital camera or fax-modem.
6. A video card and monitor (min. resolution 800×600).
7. A keyboard, mouse or other pointing device.

Installing ABBYY FineReader

The installation program will guide you through installation of ABBYY FineReader. Please close all applications prior to installing ABBYY FineReader.

To install ABBYY FineReader:

1. Insert the ABBYY FineReader 8.0 CD-ROM into the CD-ROM drive. The installation program will be launched automatically.
2. Follow the installation instructions.

If the installation program does not launch automatically:

1. Click the **Start** button on the **Taskbar** and select the **Settings/Control Panel** item.
2. Double-click the **Add/Remove Programs** icon.
3. Select the **Install/Uninstall** tab and click the **Install** button.
4. Follow the installation program instructions.

Installation options

During the installation, you will be asked to select one of the two installation options:


- **Typical** (recommended) – This option installs all components of the program, including all recognition languages. You will be prompted to choose a single interface language during installation.
- **Custom** installation – This option allows you to choose to install only specific components of the program, including all available recognition languages.

Consult the readme.htm file on the ABBYY FineReader CD-ROM if you encounter an error message.

Note: If you wish to retain your user dictionaries and patterns from a previously installed version of ABBYY FineReader, do not uninstall the older version of the program prior to installing the new version. All existing user **dictionaries and patterns** will then be available for use in the latest version.

Starting ABBYY FineReader 8.0

To start ABBYY FineReader:

- Select the ABBYY FineReader 8.0 Professional Edition (Corporate Edition) item in the **Start/Programs** menu.
- Click  on the **Microsoft Word** toolbar.
- In **Windows Explorer**, right-click the file you wish to open. In the local menu, select the **Open with ABBYY FineReader** command.

Note: Make sure your scanner is connected to your computer, plugged-in, and turned on before you start ABBYY FineReader. To install a scanner after installing the program, please consult the user guide supplied with the scanner for installation instructions. If you do not have a scanner, you can still recognize image files using ABBYY FineReader 8.0. You will find sample image files in the **ABBYY FineReader/Demo** folder on the program CD-ROM.

Installation on a Network Server or Workstation

Only the system administrator may install ABBYY FineReader 8.0 Corporate Edition on a network server. There are two stages to the installation. First, the program is installed on the server. From the server, the program can be installed on workstations using one of the four methods:

- using Active Directory
- using Microsoft Systems Management Server (SMS)
- using Task Scheduler
- from the command line

- manually in interactive mode

To install ABBYY FineReader 8.0 Corporate Edition on the server:

1. Insert the ABBYY FineReader CD-ROM into the CD-ROM drive.
2. Run **Adminsetup.exe** from the ABBYY FineReader CD-ROM.

The **System Administrator's Guide** (which can be found in the Administrator's Guide folder on the server where ABBYY FineReader is installed) provides additional information about installing ABBYY FineReader on workstations, working with the License Manager and working with the program in a local area network.

Acquiring the Image

The quality of the source image greatly affects recognition quality. In this chapter, you will learn how to scan documents for best results, how to open and read saved images (see the list of supported image formats in "Supported Image Formats" section), and how to process images to improve recognition quality (by eliminating scanning "dust" etc.).

Scanning

ABBYY FineReader communicates with the scanner through a TWAIN interface. The TWAIN standard, which was adopted in 1992, is a universal standard that unifies the interaction between a computer image input device (such as a scanner) and an external application. ABBYY FineReader communicates with a scanner through a TWAIN driver in two ways:

- through the ABBYY FineReader interface. In this case, use the **Scanner Settings** dialog and select **Use ABBYY FineReader interface**;
- through the scanner's TWAIN interface. In this case, use the scanner's TWAIN dialog to set scanning options; select **Use TWAIN-source interface**.

Each mode has its advantages and disadvantages

Using the TWAIN source interface makes the "preview image" option available so that you can set the scanning area and tune the brightness precisely, and see how these changes affect the previewed image. Every scanner has a unique TWAIN driver dialog. Consult your scanner's documentation for precise instructions on using the TWAIN dialog.

Using the ABBYY FineReader interface provides access to a couple of additional features: a) the ability to scan multiple pages with a scanner that does not have an automatic document feeder (ADF); and b) the ability to access scanning options in the batch template file (*.fnt) and use them for other batches.

Switching between modes is easy:

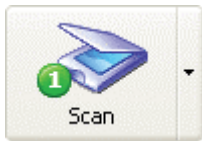
- Select the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options**), select the interface – either **Use TWAIN-Source interface** or **Use ABBYY FineReader interface**.

Note:

1. The **Use ABBYY FineReader interface** option may be unavailable (or disabled) in certain scanner models.
2. If you wish to see the **Scanner Settings** dialog in Use ABBYY FineReader interface mode, select the **Display options dialog before scanning** item on the **Scan/Open** tab (**Tools>Options**).

Important: Consult your scanner's documentation to ensure it is set up correctly. After connecting the scanner to the computer, install a TWAIN driver and/or the scanner software.

To start scanning:



Click the **1-Scan** button or select the **Scan** item in the **File** menu. The **Image** window containing a scanned image of the page will appear in ABBYY FineReader's main window.

To scan multiple pages in a row, select the **Scan multiple images** option on the **Scan/Open** tab in the **Options** dialog.

Note: To open this dialog, select the **Options...** item from the **1-Scan** button menu.

If scanning does not begin immediately, one of two dialogs will open:

- The scanner's **TWAIN-Source** dialog. In this dialog, check the scanning options and click the **OK** button (depending on your scanner model, it may be called **Done**, **Scan**, **Final**, etc.) to start scanning.
- The **Scanner Settings** dialog. In this dialog, check the scanning options and click the **OK** button to start scanning.

Tip:

To start recognition immediately after the source images are scanned, use the **Scan&Read** option:



Click the arrow at the right of the **Scan&Read** button and select the **Scan&Read** item in the local menu.

ABBYY FineReader will scan and read the images. The scanned image will appear in the Image window and the recognition results will be displayed in the **Text** window of the main window.

Setting Scanning Parameters

Recognition quality depends greatly on the quality of the scanned image. You can improve the image quality by altering the main scanning parameters: resolution, scan mode and brightness.

The main scanning parameters are:

- **Resolution** – use 300 dpi resolution for regular texts (font size 10 pts or greater) and 400–600 dpi resolution for texts set in smaller font sizes (9 pts or less).
- **Scan mode – gray.**
Scanning in grayscale mode is best for recognition purposes. During grayscale scanning, brightness is adjusted automatically.
- **Scan mode – black and white.**
Black and white scanning maximizes scanning speed but may result in the loss of some character information. This may lower recognition quality in documents of medium and low print quality.
- **Scan mode – color.**
Select this mode for documents that contain pictures, colored text or colored backgrounds, so that you can retain the original colors. In all other cases, **gray** scan mode is preferable.
- **Brightness** – a medium brightness value of around 50% should suffice in most cases. Some documents scanned in black and white mode may require additional brightness tuning.
Note: Scanning at 400 to 600 dpi resolution (instead of the default 300 dpi) or scanning in grayscale or color (instead of black & white) mode takes more time. Some scanners may take up to four times longer to scan at 600 dpi than 300 dpi.

To set scanning parameters:



- To scan images using the **ABBYY FineReader TWAIN interface**, click the **Scanner Settings** button on the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options**). The **Scanner Settings** dialog will open. Select the appropriate scanning options from the dialog.
- If you wish to scan your images using the **TWAIN-Source interface**, your scanner's TWAIN dialog will open automatically when you click the **1-Scan** button. Set the scanning parameters in the dialogue. Scanning options may have different names depending on the scanner model. For example, for brightness the word "threshold", a "sun" symbol or a black and white circle may be used. Consult your scanner documentation for a full description of available options.

Tips on Brightness Tuning

To be recognized, a scanned image must be legible. Check the legibility of the image in the **Zoom** window.

brightness – an example of an image that is appropriate for OCR

If you see that the scanned image is compromised (characters are glued or torn), consult the table below to find ways to improve image quality.

Your image looks like this:	Possible remedy:
 characters are "torn" or very light	<ul style="list-style-type: none"> ● Lower the brightness (to make the image darker). ● Scan in gray mode (to activate brightness autotuning).
 characters are distorted, glued, or filled	<ul style="list-style-type: none"> ● Increase the brightness (to make the image brighter). ● Scan in gray mode (to activate brightness autotuning).

Scanning Multi-Page Documents

ABBYY FineReader offers a specialized scanning mode (**Scan Multiple Images**) for more convenient scanning of a large amounts of pages. To enable this mode, select the **Scan multiple images** option on the **Scan/Open** tab of the **Options** dialog (menu **Tools>Options**). However, note the following:

- If you use the **ABBYY FineReader TWAIN interface**, scanning will be continuous, i.e. when one page is finished, the program will automatically start on the next.
- If you use the **TWAIN-Source interface**, the TWAIN-dialog of the scanner will remain open after scanning a page so that the next page can be placed onto the scanner immediately.

The process of scanning a large number of pages depends on whether you are using a scanner with an Automatic Document Feeder (ADF) or one without.

ADF Scanning:

1. If you are using the **ABBYY FineReader interface**, select the **Use automatic document feeder** option in the **Scanner Settings** dialog (to open this dialog, click the **Scanner Settings** button on the **Scan/Open** tab of the **Options** dialog) and the **Scan multiple images** option on the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options...**), then click **1-Scan** to start scanning.
2. If you are using the **TWAIN-Source interface**, select the **Use automatic document feeder** option in the TWAIN dialog of your scanner (remember that each scanners may name this options differently; consult your scanner documentation for details) and the **Scan multiple images** option on the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options**), then click **1-Scan** to start scanning.

Non-ADF Scanning

If you are using the **ABBYY FineReader interface**, select the **Scan multiple images** option on the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options...**) and then click **1-Scan** to start scanning.

If you are using a flatbed scanner without an ADF and the ABBYY FineReader interface, there are two ways to increase its efficiency:

- Set a pause value (i.e. the time that will elapse between the scanning of one page and the next). To do this, select the **Pause between pages** option and then set the pause value (in seconds) in the **Scanner Settings** dialog (to open this dialog, click the **Scanner Settings** button on the **Scan/Open** tab of the **Options** dialog). The scanner will pause for the predefined time before scanning the next page to allow you to place the next page onto the scanner. After the pause, scanning continues automatically.
- Select the **Stop between pages** option in the **Scanner Settings** dialog (to open this dialog, click the **Scanner Settings** button on the **Scan/Open** tab of the **Options** dialog). Each time a page scan is completed, a dialog will ask you if you wish to continue scanning. Click the **Yes** button to continue scanning or **No** to end the process.

When you have finished scanning your pages, select the **Stop Scanning** item in the **File** menu.

If you are using the **TWAIN-Source interface**:

- Select the **Scan multiple images** option on the **Scan/Open** tab in the **Options** dialog (menu **Tools>Options...**) and then click **1-Scan** to start scanning. The TWAIN dialog of your scanner will open. Click the **Scan (Final, or other)** button to start scanning.
- Scan a page, insert the next page into your scanner and click the **Scan** button in the **TWAIN dialog** of your scanner to continue scanning.
When all pages have been scanned, click the **Close** or other scanner-specific button in the TWAIN dialog of your scanner.

Tip: To have greater control over the quality of your scanned images, make sure that the **Open image during scanning** option in the **Scan/Open** group in the **Legacy Options** dialog is selected. (To open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.) This command opens each scanned page in the **Image** window immediately after it has been scanned. Reject the scanned page and halt the scanning process by clicking on **Stop Scanning** in the **File** menu. Next, re-scan the image.

Solving Scanning Problems: Your Scanner does not Support TWAIN

Even if your scanner is not TWAIN-compatible, you can still continue using ABBYY FineReader! Just do the following:

1. Create a new batch and open it. (If a batch is already open, skip this step.)
2. Set the correct recognition parameters (recognition language, document type, print type).
3. Select **Start Background Recognition** in the **Process** menu.
4. Scan the document you want to read using any image acquiring program your scanner works with. Do not close ABBYY FineReader. Save the scanned image in the folder you've saved the open ABBYY FineReader batch to; the file name should be 0001.TIF. ABBYY FineReader will pick up the image automatically and read it.
Note: If there are some pages in your batch already, the first scanned image file name should not be 0001.TIF, but XXXX.TIF, where XXXX is the number of batch pages plus one. For example, if there are 10 pages in your batch, the first scanned image file name should be 0011.tif. If you scan one more file, it should be named 0012.TIF, and so on.
5. Scan the second document and save it as 0002.TIF, etc.
6. Press **F5** to update the page list.
7. Select **Stop Background Recognition** in the **Process** menu to stop recognition.

So ABBYY FineReader will read step by step all the pages you want it to read.

Opening Images and PDF Files

You can recognize image files without using a scanner (see the list of supported image formats under "Supported Image Formats"). To open an image:

- Click on the downward-pointing arrow to the right of the **1-Scan** button and select the **Open Image** item in the local menu. An **Open** caption will replace the **Scan** caption on the button.
- Select **Open Image** from the **File** menu.
- In Windows Explorer, right-click the image file you want to open and select **Open with ABBYY FineReader** from the local menu. If ABBYY FineReader is running, the image will be added to the current batch. Otherwise, the program will be launched and the most recently used batch opened before the image is added.
- In Microsoft Outlook or Windows Explorer, click on the image file you want to open and drag it onto the minimized ABBYY FineReader window. The image will be added to the current batch and opened in the **Image** window.

Select one or several images in the **Open** dialog. The selected images will be displayed in the **Batch** window, and the last selected image displayed in the **Image** and **Zoom** windows. All selected images are copied into the batch folder. See the "General Information on Working with Batches" section for more information on batch organization and a description of how pages are displayed within batches.

Tip: If you want the opened images to be recognized right away, select the **Open&Read** mode:

1. Select the **Open&Read** item in the **Process** menu or just press **CTRL+SHIFT+D**. The **Open** dialog will open.
2. Select the images for recognition in the **Open** dialog.

Opening PDF files

The author of a PDF file can limit access to his PDF file. For example, the author may protect his file by a password or restrict certain features such as extracting text and graphics. It would be a violation of the author's copyright to access these restricted features, therefore ABBYY FineReader will ask you for a password to open such files.

Scanning Dual Pages

When scanning a bound document (i.e. a book), a dual-page scan, which scans both pages simultaneously, is easiest. You can increase recognition quality, though, by splitting the two sides after scanning, in order to perform recognition, layout analysis, and de-skewing (if necessary).

To split a dual page:

- Select the **Split Dual Pages** option on the **Scan/Open** tab (**Tools>Options** menu) prior scanning.

This command splits each dual page into two batch pages. See "General Information on Working with Batches" for more information on batches.

Note: If a dual page has been split incorrectly, deselect the **Split dual pages** checkbox and re-scan the dual page, or add the page images to the batch again. Finally, try to split the image manually using the **Split Image** dialog (**Image>Split Image**).

Adding Business Cards Images to a Batch

The most efficient way of inputting business cards is to fit as many cards as possible onto the scanner plate. After input, though, each card should be recognized as a separate page (particularly if de-skewing has been done). You may choose either automatic or manual splitting tools to separate the business card image into individual cards.


Note: This process requires that the cards be arranged in a specific order. Consult the "Working with Business Cards" section in the ABBYY FineReader Tutorial for more information.


To split the image:

1. Select the image in the **Batch** window.
2. Select **Split Image** from the **Image** menu to open the **Split Image** dialog.
3. Click on Split business cards.

Note:

1. This process removes the split page from the batch and replaces it with individual card images. For more detailed information, see the "General Information on Working with Batches" section.
2. If the image has been split incorrectly, try to split the image manually by using the **Add vertical separator/Add horizontal separator** button.
3. In order to delete all separators, click the **Remove all separators** button.

4. To move a separator, switch to **Select separator** mode (click the  button).

5. To delete a separator, switch to **Select separator** mode (click the  button) and move the separator outside the image.

Using a Digital Camera to Photograph Texts

Taking photos of documents requires some skill and practice. In this section you will learn how to set up your camera so that you can get document photos suitable for OCR. For more detailed information about the settings of your particular camera, please refer to the documentation supplied with the camera.

Before taking shots...

1. Make sure that the page fits entirely within the frame and no unwanted objects are visible.
2. Make sure that lighting is evenly distributed across the page and there are no dark areas or shadows.
3. Straighten out the page if required and position the camera parallel to the plane of the document so that the lens looks to the center of the text being photographed.

Digital Camera Requirements

Minimum Requirements

- 2-megapixel sensor
- Variable focus lens (fixed-focus cameras, common in cellphones and hand-held devices, will usually produce images unsuitable for OCR)

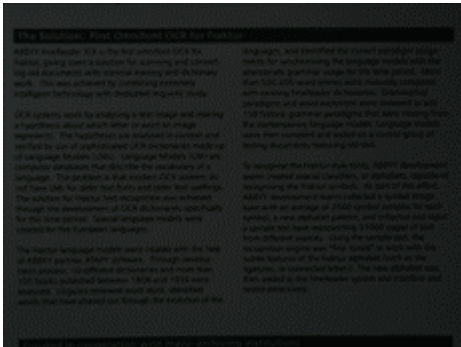
Recommended Requirements

- 5-megapixel sensor
- Flash Disable mode
- Manual aperture control or aperture priority mode
- Manual focusing
- An anti-shake system, otherwise the use of a tripod is recommended
- Optical zoom

How to Photograph Texts

Lighting

Make sure there is enough light – daylight is recommended. In the case of artificial lighting, use two light sources positioned so as to avoid shadows.



Positioning the Camera

The use of a tripod is highly recommended. The best results are obtained when shooting at the maximum optical zoom. The lens must be positioned parallel to the plane of the document and look towards the center of the text.

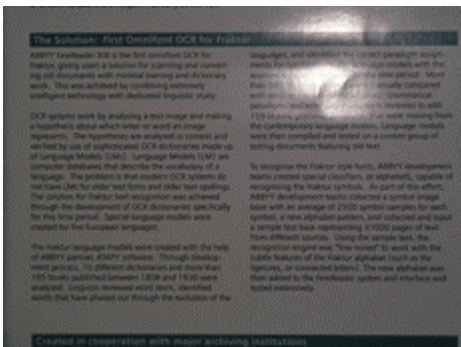
At full optical zoom, the distance between the camera and the document must be sufficient to fit the entire document into the frame. Usually this distance will be 50–60 cm.

Flash

If there is enough light, turn off the flash to avoid the glare of the page and sharp shadows. In poor lighting conditions, try using the flash from a distance of about 50 cm, but even then using additional illumination is recommended.

Important!

Using the flash for documents printed on glossy paper will cause the worst glare.



Shooting Mode

Aperture

In poor lighting conditions the recommended aperture is ~3.5 – 5.6, i.e. the maximum allowed by the camera. In bright daylight, smaller apertures will produce sharper images.

ISO Speed

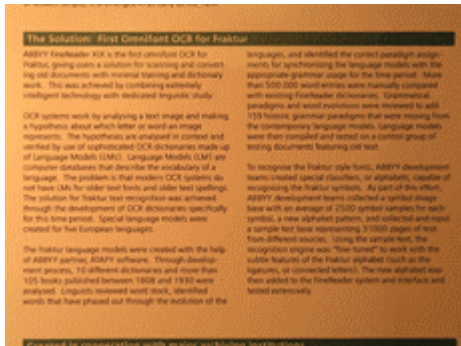
In poor lighting conditions, be sure to select a higher ISO setting.

Focus

Autofocus may not work properly in poor lighting conditions. If this is the case, focus the camera manually.

White Balance

If your camera allows, use a white sheet of paper to set white balance. Otherwise, select the white balance mode which best suits the current lighting conditions.



Additional Recommendations

Insufficient lighting will cause the camera to increase exposures, which may have an adverse effect on the sharpness of the resulting picture. Try the following:

- Enable the anti-shake system, if available.
- Use autorelease to prevent the shaking of the camera caused by pressing the shutter release button.

What do I do if...

The picture is too dark and low-contrast

Try using additional light sources. Otherwise, open up the aperture.

The picture is not sharp enough

Autofocus may not work properly in poor lighting or when trying to photograph the document from a close distance. In poor lighting conditions, try using an additional light source. When photographing a document from a close distance, try using the Macro (or Close-Up) mode. Otherwise, focus the camera manually if manual focus is supported by your camera.

If only a part of the picture is blurred, try a smaller aperture. Increase the distance between the document and the camera and use the maximum zoom. Focus on a point somewhere in between the center and a border of the image.

The flash causes a glare in the center of the picture

Turn off the flash. Otherwise, try photographing from a greater distance.

Working with an Image

Despeckle image

The recognized image may contain a large amount of "dust" (i.e. excess dots) if the original is medium-to-low print quality. Dust, when it resides close to character outlines, may adversely affect recognition quality. To decrease the number of dots:

- Select **Despeckle Image** from the **Image>Image Adjustment** menu.

To despeckle a particular block:

- Select **Despeckle Block** from the **Image>Image Adjustment** menu.


Note: Despeckling may decrease recognition quality if the original document is very faint or contains a light font. Very small characters, such as periods or commas, and parts of very thin characters may disappear.


Change image resolution

Image resolution shows the fineness of detail that can be distinguished in an image and is measured in dots per inch (dpi). ABBYY FineReader shows best OCR performance when vertical and horizontal resolution is the same and is in the range from 50 to 3200 dpi. The recommended range is 200–600 dpi, and the recommended setting is 300 dpi.

If image resolution is too small or too large, this may have an adverse effect on the quality of OCR.

Some image formats, e.g. *.bmp files, have no resolution. Sometimes an image may have non-standard resolution, e.g. 204*96 dpi, which may also adversely affect the quality of OCR.

ABBYY FineReader checks the resolution of each image and corrects if required, leaving the image dimensions unchanged. Images whose resolution has been corrected by the program are marked with  in the **Batch** window. Place the mouse cursor over such an image to see a pop-up tip.

If OCR quality for a particular image is poor, changing its resolution may help improve the quality. To change the resolution of an image marked with :

- In the **Batch** window, select the image marked with  whose resolution you wish to change. If the pop-up tip says that the image has invalid resolution, select the **Correct Resolution** command from the **Image** menu.

- In the dialog that opens, either select the type of the image (scanned image, faxed image, or screenshot) or select **Other resolution** and type in the exact resolution of the image.
- Select **Selected images** to change the resolution of the selected images. Select **All images in batch** to change the resolution of all the images in the batch. The latter option is recommended for images obtained from one and the same source.

Straighten text lines

When scanning very thick books, the text close to the binding may be distorted. Similarly, when photographing text with a digital camera, the text close to the margin may be distorted. To remedy such distortions:

- Select **Tools>Options** and click the **Scan/Open** tab. Under **Image processing**, select **Straighten text lines**.

Note: Straightening text lines may take some time.

Invert image

Some scanners invert images (turning black into white and vice versa) during scanning.



You may wish to apply the **Invert Image** option to create a uniform or standard appearance (e.g. a black font against a white background) among the documents. To do this:

- Select **Invert Image** from the **Image>Image Adjustment** menu.

Note: If you scan or open inverted images, select the **Invert image** item in the **Scan/Open** group in the **Legacy Options** dialog prior to adding these images to the batch. To open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.

Rotate or flip image

Recognition quality relies on the image having a standard orientation (the text should be read from top to bottom and all lines should be horizontal). ABBYY FineReader automatically detects page orientation during the recognition stage. If the program detects page orientation incorrectly, clear **Detect image orientation (during recognition)** on the **Scan/Open** tab and **rotate the image** manually. To do this:


- Click  or select **Rotate Clockwise** from the **Image>Rotate /Flip Image** menu to rotate the image 90° clockwise.
- Click  or select **Rotate Counter-Clockwise** from the **Image>Rotate /Flip Image** menu to rotate the image 90° counter-clockwise.
- Select **Rotate Upside Down** from the **Image>Rotate /Flip Image** menu to rotate the image 180°.

To **flip** the image:

- horizontally (around the vertical axis) – select **Flip Horizontal** from the **Image>Rotate /Flip Image** menu,
- vertically (around the horizontal axis) – select **Flip Vertical** from the **Image>Rotate /Flip Image** menu.


Clear block

You may choose to skip recognizing a particular image area or eradicate large areas of dust on the image by erasing them. To do this:

- Select  and then select the image area you want to erase by holding down the left mouse button. Release the button to erase the selected image area.

Crop image



Sometimes a scanned image may have dark borders. You can crop the unwanted black areas before running OCR. You can also use the **Crop Image** tool to reduce the image to a standard paper size, such as A4 or A5.

1. In the **Image** window, select the  tool (or select the **Crop Image** command from the **Image** menu);
2. The image will be displayed in a **Crop Image** window and its borders indicated by black lines.
 - In the drop down-list to the left, you can select the scale at which the image must be displayed in the window;
 - To crop the image, rest the mouse cursor on the color border and drag it to the desired location. Alternatively, you can rest the mouse cursor in one of the corners and drag it diagonally. The part of the image that will be removed will be displayed in gray. Click the **Crop** button;
 - To reduce the image to a standard paper size, select the desired paper size from the **Crop to** list to the right;
 - To skip cropping the current image and go to the next one, click the **Skip** button;
 - Clear the **Move to next image** box if you do not wish ABBYY FineReader to automatically move to the next image once you are finished working with the current image.

Note:

1. We recommend cropping an image before you have drawn blocks and recognized the image.
2. You can change the color of the image borders used in the **Crop** window. In order to do this, go to the **View** tab of the **Options** dialog (menu **Tools>Options**). In the **Appearance** group, select the **Crop Image Block** from the list and click the **Color** field. In the **Color** dialog, choose the desired color.

Increase/Decrease image scale

- Select  /  on the **Image** bar (from the Image window) and click on the image. The image scale will double/halve.
- Right-click the image and select **Scale**. Choose the desired scale (by percentage) from the local menu.

Get image information

You can obtain a number of parameters about your image: image width and height in pixels; vertical and horizontal resolution per inch (dpi); and image type. To do this:

- Right-click on the image and select the **Properties** item from the local menu. In the dialog that opens, select the **Image** tab.

Print image

You can print the image in the Image window, the pages selected in the Batch window, or all batch page images. To do this:

- Select **File>Print>Image**. The **Print** dialog will open. Set the desired printing parameters (the printer to be used, number of pages to be printed, the number of copies etc.).

Undo the last action

- Click the **Undo** button on the **Standard** bar .

Tip: To reverse an **Undo** action, click the **Redo** button on the **Standard** bar .

Page Numbering

A number is assigned to each scanned page. The default number is the number of the last batch page plus one.

You may set page numbers manually if you want to retain the original page numbers in the document or if you want to scan pages according to page number.

To specify page numbers:

- Select **Ask for page number before adding page to the batch** on the **Scan/Open** tab (**Tools>Options** menu).

To scan a large number of double-sided pages according to page number:

1. Select **Ask for page number before adding page to the batch** on the **Scan/Open** tab (**Tools>Options**).
2. Specify a number for the first scanned page in the **Page number** dialog, then select **Odd and even separately** in the **Page numbering** field. Select an order for the pages: ascending or descending to reflect the way in which the double-sided pages have been entered into the automatic document feeder (i.e. whether the last page or the first page has been placed on top).

Batch Image Options

Select **Convert color and gray images to black and white** to scan images in grayscale using the **TWAIN-Source interface**. The scanned images will not retain color pictures or colored fonts or backgrounds. This option reduces the amount of disk space needed to store scanned images.

Note: This option can be found in the **Legacy Options** dialog. To open this dialog box, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.

Page Layout Analysis

Before starting the recognition process, ABBYY FineReader must know which image areas it needs to recognize. To achieve this, the page layout analysis process identifies text blocks, picture blocks, table blocks, and barcode blocks.

In this chapter you will learn more about: when manual page analysis is necessary; what block types are available; how to edit blocks drawn using automatic layout analysis; and how to streamline the layout analysis with block templates.

General Information on Page Layout Analysis

Page layout analysis can be done either automatically or manually. In most cases, ABBYY FineReader manages the complex task of analyzing page layout by itself. Start automatic analysis by clicking on the **2-Read** button. Recognition and layout analysis are performed simultaneously.



Click this button to start the process of reading an open image. To change the button mode, click the arrow at the right of it and select the necessary item in the local menu.

Note: Stand-alone page layout analysis is also available (**Process>Read>Analyze Layout** menu). This process may be needed at times, but often this approach provides inferior page layout analysis, since coupled layout analysis/recognition uses information acquired during recognition to improve layout analysis.

You may opt to draw blocks manually if:

1. Only a part of a page needs to be recognized;
2. Automatic layout analysis has drawn blocks incorrectly.

Tip:

- In some cases, the quality of the automatic layout analysis can be improved by changing the page layout analysis options. To view the current layout analysis options, go to the **Read** tab, **Tools>Options** menu.
- If the application has drawn some blocks incorrectly, it is often faster to edit the incorrect blocks with the block editing tools than to delete the blocks and draw them again manually.

Block Types

Blocks are image areas enclosed in frames. Blocks tell the system which image areas should be recognized and in what order. The blocks also influence how the original page layout is retained. The differently colored frames indicate different types of blocks. The frame colors of the blocks can be changed on the **View** tab of the **Options** dialog (**Tools>Options** menu) in the **Appearance** group. Select the required block type in the **Item** field and the desired color in the **Color** field.

The following block types are available:

Recognition Area – this is used for automatic recognition and analysis. After the **2-Read** button is clicked, all blocks of this type will be automatically analyzed and recognized.

Text – this is used for text image areas and should only contain single-column text. If there are pictures within the text, draw separate blocks around them.

Table – this is used for table image areas or for areas of text that are structured in a table. When the application reads this type of block, it draws vertical and horizontal separators inside the block to form a table. This block is represented as a table in the output text. You can draw and edit tables manually.

Picture – this is used for image areas that contain pictures. This type of block may enclose an actual picture or any other object that should be displayed as a picture (e.g. a section of text).

Barcode – this is used for barcode image areas. If your document contains a barcode that should be displayed as a series of numbers and letters rather than as a picture, draw a separate block for the barcode and set the block type to barcode.

Note: If you wish ABBYY FineReader to read barcodes on your documents automatically, make sure that the **Look for barcodes** option is selected in the **Read** group in the **Legacy Options** dialog, otherwise clear this option. (To open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.)

Barcode types

Barcode types	Code 3 of 9
	Check Code 3 of 9
	Code 3 of 9 without asterisk
	Codabar
	Code 93
	Code 128
	EAN 8
	EAN 13
	IATA 2 of 5
	Inerleaved 2 of 5
	Check Inerleaved 2 of 5
	Matrix 2 of 5
	Postnet
	Industrial 2 of 5
	UCC-128
	UPC-A
	UPC-E
	PDF417

Automatic Page Layout Analysis Options

As a part of automatic page layout analysis the following types of blocks are drawn: text, table, picture, and barcode.

To start automatic layout analysis (and text recognition), click the **2-Read** button. Before clicking this button, however, select the table analysis options.



Click this button to start the recognition of an open image. To change the button mode click the arrow at the right of it and select the necessary item in the local menu.

Table analysis options

Usually, the application divides tables into rows and columns automatically. If additional tuning of table options is needed, open the **Legacy Options** dialog and in the **Read** group select the desired item. (To open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog). Change these options if:

- automatic page layout analysis has drawn the table rows and columns incorrectly;
- the document contains a large number of simple tables of the same type (i.e. there are no merged cells or there is always only one line of text per cell).

1. Use the **One line of text per cell in table** option if your table has no (or minimal) black separators and each cell has only a single line of text. For example:

Kilometers	Miles
1	0.62
5	3.2

– this table has only one line of text per cell

Physical phenomenon	t, degrees centigrade
Water boiling point	100
Water freezing point	0

– this table has more than one line of text per cell

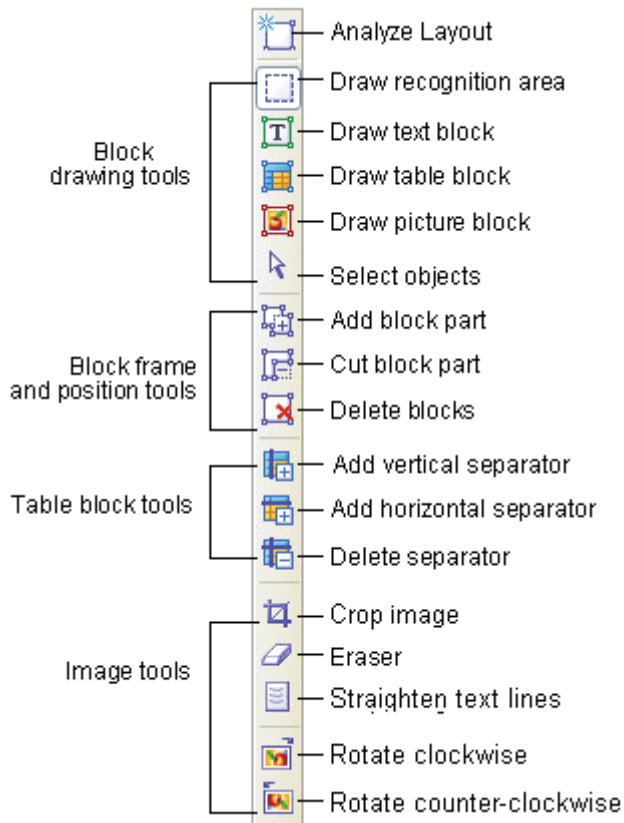
2. Use the **No merged cells in table** option if your table has no merged cells in it. For example:

Temperature	
Degrees centigrade	Degrees Kelvin
-273	0
100	373

– the **Temperature** cell is a merged cell

Note: Do not select **One line of text per cell in table** and/or **No merged cells in table** if the text contains tables with differing structures. Selecting these options may result in errors during layout analysis and may adversely affect recognition quality.

Drawing and Editing Blocks Manually



To create a new block:

1. Select one of the following tools:



to draw a recognition area;



to draw a text block;



to draw a picture block;



to draw a table block.

2. Position the mouse at the point where you want a corner of your block to be. Hold down the left mouse button and drag the mouse pointer to the point where you want the opposite block corner to be.
3. Release the mouse button.

A frame will enclose the selected image area.

You may change the drawn block type to any of the following: **Recognition Area**, **Text**, **Table**, **Picture**, or **Barcode**. To change a block type:

- Right-click the block and select the **Change Block Type** item followed by the corresponding block type in the local menu.


Modifying blocks

To move block borders:


1. Click the block border and hold down the left mouse button. The mouse pointer will become a two-headed arrow.
2. Drag the pointer in the desired direction.
3. Release the mouse button.

Note: If you click a block corner, you can move both horizontal and vertical borders of the block simultaneously.

To add a rectangular block part:

1. Select the  tool.
2. Click the block you wish to add to. Press and hold down the left mouse button and drag the mouse pointer diagonally. Select the desired image area and release the button. The resulting rectangle will be added to the block.
3. If necessary, move the block border.

To cut out a portion of a rectangular block:


1. Select the  tool.
2. Click on the portion of the block you wish to cut. Press and hold down the left mouse button then drag the mouse pointer diagonally. Select the desired area and release the button. The selected rectangle will be cut from the block.

3. If necessary, move the block border.









Note:

1. You can alter block borders by adding new nodes (splitting points). Use the mouse to move split border segments in any direction.
To add a new node, press **SHIFT**, place the mouse pointer to where you want a new node (the pointer will become a cross) and click on the border. A new node will be created.
2. ABBYY FineReader imposes certain limitations on block form. To be successfully recognized, text lines within blocks must be unbroken. To enforce these requirements, ABBYY FineReader automatically corrects block borders as parts are added or deleted. For example, if you delete a portion from the top or bottom of a block, a whole block corner will automatically be cut. Similarly, if you try to cut off a part from between the two upper or lower corners, the application will cut the right block corner (upper or lower) as well. The program will also forbid operations that involve moving the segments that form the block borders.

To select a block or a group of blocks:

- Select the  tool and click on the desired block or press the left mouse button and draw a rectangle around all the blocks you want to select.

Note: You can select one or more blocks using the block drawing tools. To select several blocks at once hold down **SHIFT** or **CTRL**

with one of the following tools activated: , ,  or . Drag the arrow over the blocks you want to select. To invert the selection (i.e. to select an unselected block or vice versa), hold down the **CTRL** key while one of the following tools is activated: , ,  or  and drag the arrow over the desired blocks.

To move blocks:


- Hold down **ALT** with one of the following tools activated: , , ,  or  and move the blocks.

To renumber blocks:

1. Select the  tool.
2. Click the blocks in the desired order. The contents of blocks will be displayed in the output text in the same order.

Note: If you renumber blocks on a previously recognized image, the recognized text in the draft mode of **Text** window will be re-arranged to reflect the new numbering.

To delete a block:

- Select the  tool and click the block you wish to delete, or
- Select the blocks you wish to delete and press **DEL** on the keyboard.

Note: If you delete a previously recognized block, its associated text in the **Text** window will be deleted as well.

To delete all image blocks:

- Select the Delete ALL Blocks and Text item in the Image menu.

Note: If you delete blocks on an image that has already been recognized, the recognized text in the **Text** window will also be deleted.

Editing a table

To edit a table, select one of the following tools on the **Image** toolbar:



– to add a vertical separator;



– to add a horizontal separator;



– to remove a separator.

To merge several cells:

- Select the **Merge Cells** item in the **Image>Table Cells** menu.

To split previously merged cells:

- Select the **Split Cells** item in the **Image>Table Cells** menu.

To merge table rows (the division into columns is retained)

- Select the **Merge Rows** item in the **Image>Table Cells** menu.

Manual Table Layout Analysis

Tip: If automatic table layout analysis has incorrectly drawn table rows and columns, editing the automatic analysis results instead of deleting all the blocks and re-drawing them manually is usually more efficient.

Editing a table manually:

Use the following **Image** toolbar tools to edit a table:



– Add a vertical separator



– Add a horizontal separator



– Remove a separator

If the table cell only contains a picture, select the **Treat cell as picture** item in the **Block Properties** dialog (menu **View>Properties**). If the table cell contains both text and pictures, draw a separate picture block (or blocks) inside the cell.

To merge table cells or rows:

- Select the Merge Cells or Merge Rows item in the Image>Table Cells menu.

Note: You can split previously merged cells using the **Split Cells** command (**Image>Table Cells** menu). The **Merge Rows** option does not affect the division of the table into columns.

Note: To avoid drawing horizontal and vertical separators manually, draw a separate table block, then right-click on it. Select **Analyze Table Structure** in the local menu. The system will then draw all the necessary separators. Should the system draw any separators incorrectly, you can edit the table manually.

Using Block Templates

If you are processing a large number of documents with an identical layout (e.g. forms or questionnaires), analyzing the layout of every page will be time consuming. To save time, create a block template (i.e. a standard set of blocks of a particular type that correspond to a page's layout) and then apply the template to all pages you wish recognized that have the same layout.

Note: Documents should always be scanned using their respective template(s) and using the resolution that was used to create the template(s).

To create a block template:

1. Open an image and draw the blocks automatically or manually.
2. Select **Save Blocks Template** in the **Image** menu. The **Save Blocks Template As** dialog will open. Type a file name for the block template in the dialog.

To load a block template:

1. Click the **Batch Window** and select the pages you wish to apply the block template to.
2. Select **Load Blocks Template** in the Image menu. The **Open Blocks Template** dialog will open.
3. Select the relevant block template file in the dialog.
4. Click the appropriate **Apply to** item in the group. **All pages** applies the block template to all batch pages, while **Selected pages** applies the block template only to selected pages.
5. Click the **Open** button.

Recognition

The aim of OCR is to read text from a source image and retain the source page layout. Before this can be done, however, the main recognition parameter – recognition language – needs to be set. This chapter deals with the recognition parameters and other important recognition issues, including the use of different recognition settings etc.

General Information on Recognition

Note: Always ensure that the following options have been correctly set before you start recognition: recognition language, recognition mode and source text print type.

You may:

1. Recognize a block or several blocks drawn on an image.
2. Recognize an open page or all pages selected in the **Batch** window.
3. Recognize all unrecognized batch pages.
4. Recognize all pages in background mode. Background mode allows you to edit and recognize pages at the same time.
5. Recognize pages in training mode. Training mode is used for recognizing texts set in decorative fonts or for processing large volumes (more than a hundred pages) of documents of inferior print quality.
6. Recognize the same batch on several workstations.

Note: This feature is available only in ABBYY FineReader 8.0 Corporate Edition.

To start recognition:

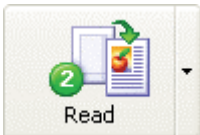
- Either click the **2-Read** button on the **WizardBar** toolbar, or
- Select the necessary item in the **Process>Read** menu:

Read – to recognize the open page or all the pages selected in the **Batch** window;

Read All Pages – to recognize all unrecognized batch pages;

Read Block – to recognize a block or several blocks drawn on the image;

- Select **Process>Start Background Recognition** to start recognition in background mode.

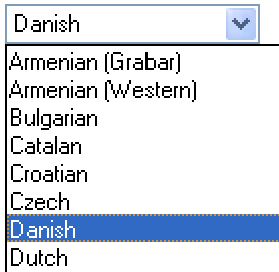


By default, the **2-Read** button recognizes the open image. To change button mode, click the arrow to the right of the button and select the mode of your choice in the local menu.

Note: When you perform OCR on a page that has already been recognized, recognition will only be carried out on new or modified blocks.

Recognition Languages

ABBYY FineReader recognizes both mono- and multilingual (e.g. English and French) documents. To set the text recognition language, select it in the drop-down list on the **Standard** toolbar.



To recognize a multi-lingual document:

1. Select the **Select multiple languages** item in the language list on the **Standard** toolbar. The **Recognition language** dialog will open.
2. Select the language of your choice in the **Recognition language** dialog.

Note:

1. If you find that you often use a certain language combination, you can create a new language group that includes the languages you most often use.
2. Increasing the number of the recognition languages used simultaneously may have an adverse effect on recognition quality. A reasonable number of languages to use simultaneously is 2–3.
3. Before recognizing a document, ensure that the fonts selected on the **Save** tab support all the characters contained in the recognition language(s) chosen, otherwise the recognized text will be displayed incorrectly ("?" or " " symbols will appear instead of letters).

You may find that your chosen recognition language is not listed. This can be because of one of the following reasons:

1. The language is not supported by ABBYY FineReader. See the complete list of recognition languages under "Supported Languages".
2. The language hasn't been included in the recognition language list displayed on the **Recognition** toolbar. To add a language select the **Choose more languages** item in the language list on the **Standard** toolbar. The **Recognition language** dialog will open. Select the language of your choice in the dialog.
3. The language was disabled during custom installation.
Note: Always ensure that you use the same folder as the one that contains ABBYY FineReader.
4. Your copy of ABBYY FineReader was purchased in an online store. To reduce download times, the distribution package sold via online stores includes only the most common interface and recognition languages. To download more languages, select **Start/Programs/ABBYY FineReader 8.0/Download more languages** and follow the instructions of the program.

To show/hide a language in the drop-down list on the toolbar:

- Select the language of your choice in the Language Editor dialog (Tools>Language Editor) and then check or uncheck the Show this language in the drop-down list on the toolbar item.

Tip: It is even possible to set a recognition language for an individual block. To do this, right-click the block and select the **Properties** item in the local menu. The **Properties** dialog will open. Select the **Block** tab in the dialog and then select the block recognition language in the **Languages** field on the tab.

Source Text Print Type

As a rule, source text print type is determined automatically. To ensure that this is the case, select **Autodetect** in the **Print Type** group in the **Legacy Options** dialog (to open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog).

When recognizing draft mode dot matrix printouts or typewritten texts, recognition quality can sometimes be increased by selecting another **print type**:

- Select the **Typewriter** item if you wish to recognize typewritten texts.
- Select the **Dot matrix printer** item if you wish to recognize dot matrix printouts.

software An example of draft mode dot matrix text. Character lines are made up of individual dots.

software An example of typewritten text. All letters are of equal width (compare, for example, "w" and "a").

To change print type:

- Select the print type of your choice in the **Print Type** group in the **Legacy Options** dialog (to open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog).

Note: Once you have completed recognition of typewritten texts or dot matrix printouts, remember to re-enable the **Autodetect** item to recognize normal texts once again.

Other Recognition Options

Recognition mode

ABBYY FineReader 8.0 allows you to choose speed or quality during the recognition process. Two recognition modes are available:

- **Thorough**
In this mode, ABBYY FineReader will analyze and recognize both simple documents and documents with complex layouts, in particular documents with text on a color or raster background and documents with complex tables (including tables with white grid lines and tables with color cells).
Note: As compared to the **Fast** mode, the **Thorough** mode consumes more time but ensures better recognition quality.
- **Fast**
This mode is recommended for processing large amounts of documents with simple layouts and good image quality.

You can set up the recognition mode on the the **Read** tab in the **Options** dialog (menu **Tools>Options**).

PDF recognition

ABBYY FineReader 8.0 extracts text data from PDF files and uses these data for recognizing PDF documents. Text data extraction speeds up the recognition process up to 2 – 3 times. However, PDF files may have non-standard encoding. In this case text content can only be recovered with OCR. If you are not satisfied with the recognition quality of a PDF document:

- On the **Read** tab in the **Options** dialog (menu **Tools>Options**), select the **Recognize PDF files as images** option in the **PDF recognition** group and re-read the document.

Barcode recognition

If you want ABBYY FineReader to recognize barcodes on your documents automatically, make sure that the **Look for barcode** option is selected in the **Read** group of the **Legacy Options** dialog. ABBYY FineReader will create separate barcode type blocks for them; barcodes will be displayed as a series of letters and numbers in the recognized text. The full list of the barcode types supported by ABBYY FineReader 8.0 see in "Block Types".

Note: To open the **Legacy Options** dialog, click the **Legacy Options** button on the **General** tab in the **Options** dialog (**Tools>Options**).

Show image during recognition

When processing large numbers of pages, recognition is invariably faster if the processed image is not displayed on-screen. To run recognition without displaying the image:

- Clear the **Show image during recognition** item in the **Legacy Options** dialog (to open the **Legacy Options** dialog, click the **Legacy Options...** button on the **General** tab in the **Options** dialog).

Text orientation

If the application incorrectly recognizes blocks containing vertical text (a text block or a table cell):

- Right-click on the block with vertical text and select the **Properties** item in the local menu. The **Block properties** dialog will open. Select the relevant item in the **Text orientation** list in the dialog and re-recognize the image.



Inverted or flipped block

If the application incorrectly recognizes blocks containing inverted or flipped text (a text block, a table cell, or a whole table):

- Right-click on the block concerned and select the **Properties** item in the local menu. The **Block properties** dialog will open. Select the **Inverted** or **Flipped** item in the dialog and re-recognize the image.

Background Recognition

If you wish to edit previously recognized pages and run recognition at the same time, you may find background recognition mode useful. To start background recognition:

- Select the **Start Background Recognition** item in the **Process** menu.
The sign  will appear in the status line at the bottom of ABBYY FineReader's main window. If **Details** view mode is active in the **Batch** window (to activate **Details** view mode, right-click on the **Batch** window and select **Batch Window>Details** in the local menu), the page currently being recognized will have the  icon displayed in the **Opened by** column.

When background recognition mode is activated, recognition will resume automatically if an unrecognized page is added to the batch.

Note: Running Background mode in the case of multiprocessor systems leads to an increase in recognition speed if the batch being processed contains a large number of pages.

To stop Background Recognition:

- Select the **Stop Background Recognition** item in the **Process** menu.

Note: Background recognition mode uses recognition options active at the moment it was started.

Recognition with Training

As previously mentioned, ABBYY FineReader can read texts set in practically any font regardless of print quality. Consequently, no prior training is normally required before recognition can take place. ABBYY FineReader, nevertheless, features a number of user pattern training tools.

Train User Pattern mode may come in useful when:

1. recognizing texts set in decorative fonts;
2. recognizing texts containing unusual characters (e.g. mathematical symbols);
3. recognizing large volumes (more than a hundred pages) of texts of low print quality.

Tip: Use **Train User Pattern** mode **only if one of the above applies**. In other cases you may obtain a slight increase in recognition quality, but the time and effort involved will probably outweigh the benefit.

Pattern training works as follows. One or two pages are recognized in training mode, and, subsequently, a pattern is created. ABBYY FineReader then uses this pattern to aid recognition of the remaining text.

Sometimes two or even three characters may get "glued" together, and ABBYY FineReader may be unable to enclose each character in an individual frame to separate them. If this is the case (i.e. you cannot move the frame so that it contains only one whole character and no other character parts), you can train ABBYY FineReader to recognize the whole inseparable character combinations. Examples of character combinations frequently found glued together include ff, fi, and fl. Such combinations are referred to as ligatures.

Note:

1. A pattern is only useful in the case of documents that have the same font, font size, and resolution as the document used to create the user pattern.
2. Each pattern is created for a particular batch. Consequently, if a batch is deleted, its user pattern is also deleted. Patterns can, however, be copied into other batches. To transfer a user pattern to another batch, simply save the batch options in a batch template format file.
3. If you switch to recognizing texts set in a different font, always disable any user patterns – choose the **Do not use user pattern** item on the **Read** tab, menu **Tools>Options**.

To train a user pattern:

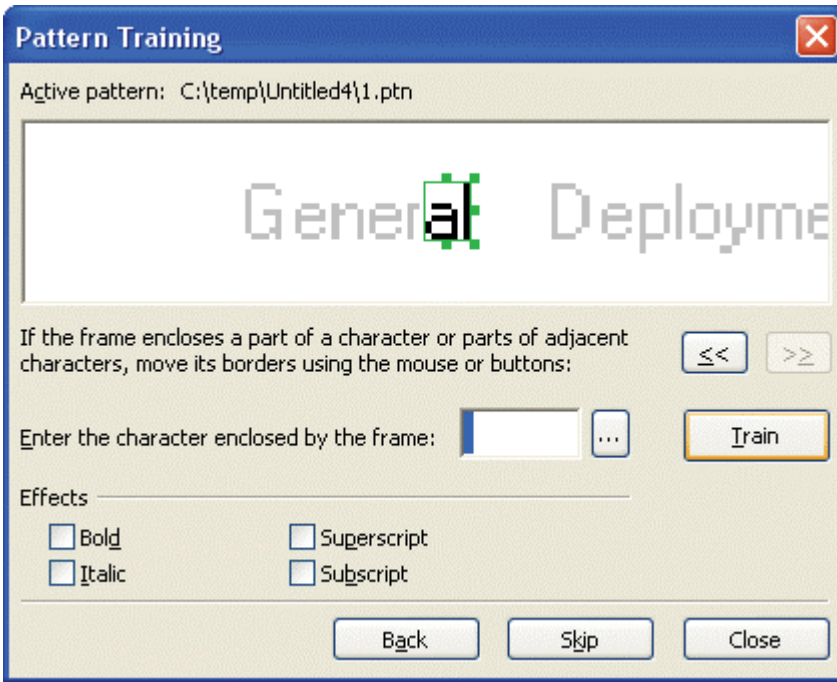
1. Start **Train user pattern** mode – click the **Train user pattern** radio button on the **Read** tab, **Tools>Options** menu, in the **Training** group. The default pattern name ("Default") will be displayed in the status line.
2. Click the **2-Read** button.
3. Train your pattern – recognize one or more pages in **Train user pattern** mode. Trained characters are saved in the default pattern. Once you have completed training the pattern, ABBYY FineReader will save the pattern (**Default.ptn**) in the current batch folder.
4. Edit your pattern.
5. Deactivate training mode (click the **Use user pattern** button on the **Read** tab).
6. Recognize the rest of the text – click the **2-Read** button.

Note:



1. To create several patterns for the same batch, use the **Pattern Editor** dialog (click the **Pattern Editor** button on the **Read** tab or select the **Tools>Pattern Editor** menu item). Create a new pattern (click the **New** button in the dialog) and select it (click the **Set Active** button). Working with a created pattern is no different to working with a default pattern (see steps 1–5). Keep in mind, however, that only one pattern may be active at any one time.
2. If you've created several patterns for the same batch, the active one will be the pattern that was last created. The active pattern name is displayed in the status bar. To activate another pattern, select the pattern of your choice in the pattern list in the **Pattern Editor** dialog (**Tools>Pattern Editor** menu) and click the **Set Active** button. Then click the **Use user pattern** button on the **Read** tab, **Tools>Options** menu, in the **Training** group.
3. If the **Use built-in patterns** option is set, ABBYY FineReader will read all texts using its built-in patterns and stop only at uncertain characters. If you are training the system to read decorative and/or non-standard fonts (for example, Tibetan) the use of built-in patterns may result in characters being read incorrectly. If this occurs, disable the use of built-in patterns (clear the **Use built-in patterns** checkbox on the **Read** tab) and train the system to recognize each unknown character it is likely to encounter.

How to Train a User Pattern


1. Make sure the **Train user pattern** radio button on the **Read** tab (**Tools>Options** menu) in the **Training** group is enabled.
2. Click the **2-Read** button. ABBYY FineReader will start recognition. Whenever it comes across an unknown character, the **Pattern Training** dialog will open, and the character image will be displayed within it.



Training to recognize a character:

The frame in the top dialog window should enclose **a single character**, and this character **must be fully enclosed** by the frame. If the frame encloses only part of a character or more than one character, click the frame borders and move them so that the above-stated requirements are met. The  and  buttons move the frame border as well (and are useful for training italic symbols – see below). Once you have positioned the frame correctly, type in the character and click the **Train** button.



Note:

1. You may only train the system to read characters included in the alphabet. If you wish to train ABBYY FineReader to read characters that cannot be entered from the keyboard, use a combination of two characters to denote these non-existent characters or copy the required character from the **Character Table** (click the  button in the **Pattern Training** dialog to open the **Character Table**).
2. If you wish to train the system to retain character formatting, select the corresponding **Italic** or **Bold** item in the **Pattern Training** dialog before clicking the **Train** button.
3. Make sure that only uppercase/lowercase characters are entered when training uppercase/lowercase character images respectively.

If you make a mistake during training, click the **Back** button to return the frame to its previous position. The last "image-character" pair to be entered will automatically be removed from the pattern. Note that this "undo" function is limited to the last word trained.

Training to recognize ligatures

A ligature is a combination of two or three "glued" characters, for example, fi, fl, ffi, etc. These characters are difficult to separate because they are "glued" during printing. In fact, better results can be obtained by treating them as single compound characters. Training ligatures is the same as training separate characters:

1. Type the necessary character combination and click the **Train** button.
2. The frame in the top dialog window should enclose **the entire ligature**. You can move the frame border using the mouse or by clicking the  and  buttons.

Each pattern may contain up to 1000 new characters. However, you should not create too many ligatures, as it may affect the recognition quality negatively.

You should also take the following limitations into account when you train ABBYY FineReader:

1. ABBYY FineReader doesn't differentiate between certain characters which are usually considered different. Such images are recognized as one and the same character. For example, the straight (') , right (') and left (') apostrophes are kept in the pattern as one character – the straight apostrophe. Thus, you'll never see the right and left apostrophes in the recognized text, even if you try to train them.
2. In some cases a certain image is recognized as a certain character depending on its environment.

How to Edit a User Pattern

You may find useful to edit your newly created pattern before you start recognition. An incorrectly trained pattern will affect the recognition quality negatively, so it is useful to check your pattern.

The pattern should contain only whole characters or ligatures. Characters with cut edges and incorrectly labeled characters should be removed from the pattern.

To edit a user pattern:

1. Select the **Pattern Editor** item in the **Tools** menu. The **Pattern Editor** dialog will open.
2. Select the necessary pattern and click the **Edit** button in the dialog. The **User Pattern** dialog will open.
3. Select a character and click the **Properties** button to edit the character caption and set the correct typeface: italic, bold, subscript or superscript. You may also click the **Delete** button to remove incorrectly trained characters from the batch.

User Languages and Language Groups

In addition to the built-in languages and language groups, you may also create a new language or make up a new language group (from languages supported by ABBYY FineReader) and use it for recognition.

You may want to create a new language if you need:

1. To connect a user dictionary.
 - For example, you need to recognize an English text containing many abbreviations. You may create an abbreviation dictionary, create a new language and connect the dictionary to the language. Then you may make up a new language group consisting of the English language (with the system dictionary) and your new language (with the abbreviations dictionary) and use this language group to recognize your texts.
2. To recognize documents of certain special types, for example:
 - supermarket product lists where there are only product codes. Such codes usually contain only numbers and probably several letters. You may then create a new language consisting of the necessary characters only and use it for recognizing these documents.
 - documents set in capitals only. You may increase the recognition quality if you create a language in which all lowercase letters are prohibited.

You should create a language group if you use a particular language combination often.

To create a new language or language group, open the **Language Editor** dialog (**Tools** menu, **Language Editor** item).


How to Create a New Language

To create a new recognition language:

1. Select the **Language Editor** item in the **Tools** menu.
2. Click the **New** button and in the resulting dialog select the **Create a Copy of the Language** button, then select your preferred source language.
3. The **Simple Language Properties** dialog will open.



Set the following language parameters for the new language (all parameters are entered in the Simple Language Properties dialog):

1. The new language name.
2. The basic alphabet to be used by the new language. This parameter is set in the **Alphabet** field. If necessary, edit the alphabet by clicking the  button.
3. The dictionary to be used by the application (for both recognition and spelling check purposes). You may choose one of the following:
 - **None** (no dictionary to be used)
 - **Built-in** (the dictionary supplied with ABBYY FineReader)
 - **User dictionary**

To add words to the dictionary or to use an existing user dictionary or text file in Windows (ANSI) or Unicode encoding (the only requirement is that words be separated by spaces or other non-alphabetic characters) click the **Edit Dictionary** button.

Note: The spelling checker will consider user dictionary words to be correct if they are found in the text in one of the following capitalizations: dictionary set capitalization; lowercase only; uppercase only; first letter – capital, remaining letters small. Examples include:

Dictionary set capitalization:	Correct occurrences of the word:
abc	abc, Abc, ABC
Abc	abc, Abc, ABC
ABC	abc, Abc, ABC
aBc	aBc, abc, Abc, ABC

- **Regular expression** (used to specify the grammatical rules of the new language; see Regular Expressions for details).

Note:

1. Click on the **Advanced** button in the **Simple Language Properties** dialog to set advanced properties for the new language, e.g. characters to be ignored, prohibited characters, etc.
2. By default, all new user languages are saved into the batch folder. Note that ABBYY FineReader Corporate Edition allows you to specify the folder to which the language should be saved. For more information on group work with user languages and dictionaries, see "Group work with the same user languages and user dictionaries".

How to Create a New Language Group

If you often recognize texts written in a certain language combination, say, English–German, you may create a language group combining these languages. The created group will be displayed in the language list on the **Standard** toolbar.

Note: You can specify the recognition languages to be used in the language list on the **Standard** toolbar. To do this, select the **Select Multiple languages** item in the list. The **Recognition Language** dialog will open. Select the languages you need in the dialog.

To create a recognition language group:

1. Select the **Language Editor** item in the **Tools** menu and click the **New** button. A dialog will open. Select the **Create a new group of languages** item in the dialog.
2. The **Language Group Properties** dialog will open.



Set the following new language group parameters (all parameters are set in the Language Group Properties dialog):

1. Group name.
2. Languages contained in the group.

Note:

1. If you know that your text will not contain certain characters, you may wish to specify these so-called prohibited characters in the relevant language group's properties. Specifying such characters can increase both recognition speed and quality. To specify prohibited characters, click the **Advanced** button in the **Language Group Properties** dialog. The **Advanced Language Group Properties** dialog will open. Specify the set of prohibited characters in the **Prohibited characters** line.

2. By default, the newly created user language group will be saved in the batch folder. In the case of ABBYY FineReader Corporate Edition, you can specify the destination folder. For more information on group work with user languages and dictionaries, see "Group work with the same user languages and user dictionaries".

Checking and Editing Text

Once recognition is over, you will see the recognized text displayed in the **Text** window. The **Text** window is ABBYY FineReader's built-in editor, used to check recognition results and edit any recognized text.

The ABBYY FineReader text editor has two distinctive features:

1. A built-in spelling check system (see the list of languages with spell check support under "Supported Languages").
2. A convenient visual aid: the source image of the text line being edited is displayed in the **Zoom** window.

The built-in spelling checker features:

1. Tools for finding uncertain words (words containing uncertain characters).
2. Tools for finding misspelt words.
3. Tools for adding unknown words to the ABBYY FineReader dictionary. Adding words to the dictionary improves recognition quality.

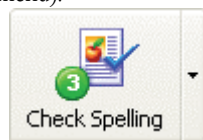
Checking Text in ABBYY FineReader

Uncertainly recognized characters and words not found in dictionary are highlighted in different colors. By default, light blue is used for uncertain characters and pink for words not found in the dictionary. To change the colors used:

- Select the **Uncertain Character** (or **Not in Dictionary word**) item followed by the color of your choice in the **Color** item on the **View** tab (**Tools>Options** menu) in the **Appearance** group.

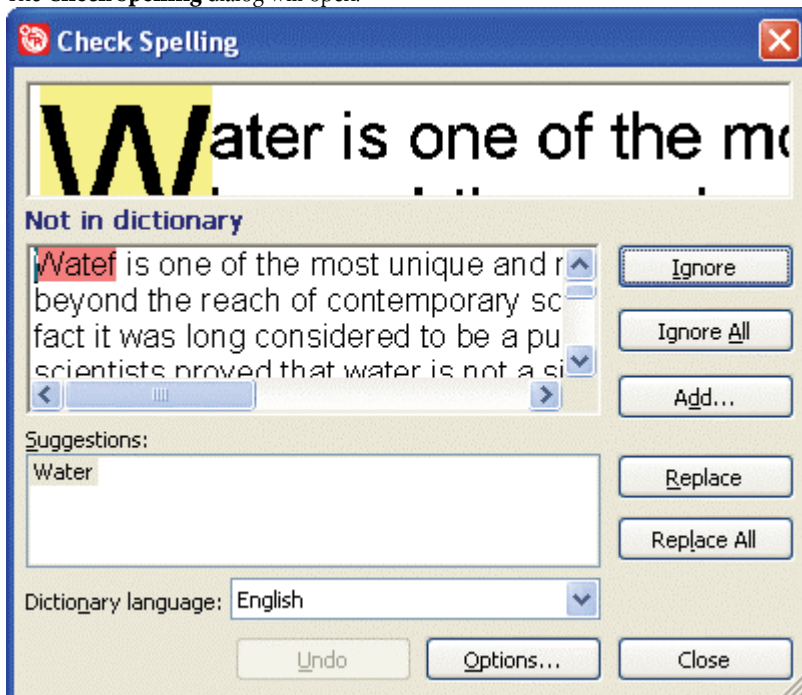
To check recognition results:

1. Click the **3-Check Spelling** button on the **WizardBar** toolbar (or select the **Check Spelling** item in the **Tools** menu).



The spelling checker looks for uncertain words, words not found in dictionaries and misspelled words. To check spelling, click this button or press **F7**.

2. The **Check Spelling** dialog will open.



3. There are three windows in the **Check Spelling** dialog. The top window is similar to the ABBYY FineReader **Zoom** window and displays the original image of the word. The middle window displays the word itself, and the line above displays the name of the error type. The **Suggestions** window at the bottom provides you with replacement suggestions (if any exist). Note that suggestions are based on the dictionary selected in the **Dictionary language** drop-down list; any language may be chosen from this list.
Note: You can enlarge the **Check Spelling** dialog to make it easier to check and edit text. Simply click the dialog border; the mouse pointer will become a double-headed arrow. Drag the border to make the dialog larger or smaller.
4. If words have been misspelt, you can do one of the following:

- Click the **Ignore** button to leave the word unchanged.
- Click the **Ignore All** button to leave all such words in the text unchanged.
Note. When you click the **Ignore** or **Ignore All** button, the "uncertain" flag is removed from the word, i.e. the system assumes that the word no longer contains any unrecognized or uncertain characters and no longer needs to be highlighted. As a result, when you export such words in PDF format and select the **Replace uncertain words with images** mode, the words for which the "uncertain" flag has been removed will not be replaced with images.
- Select a replacement suggestion and then click the **Replace** or **Replace All** button to replace the current word or all such words in the text. If no correct suggestion has been made for the word in the **Suggestions** window, you can enter one yourself in the middle window. (**Important:** when you switch to edit mode, certain buttons may change function and adopt new captions.) Click the **Confirm** (**Confirm All**) button to change the current word (or all such words) in the text and move to the next uncertainly recognized word.
- Click **Add...** to add the word to the dictionary. Once a word is added, the application will consider all subsequent occurrences of this word in any of its word forms to be correct.
- Click **Options...** to set the spelling check options.
- Click **Close** to close the dialog window.

Moving between uncertain words

To check the recognition results quickly, you can use the  button and  button to move to the next or previous uncertain word, respectively.

You can also use the **F4 (SHIFT F4)** hotkey to navigate between uncertain words.

Check and Edit Text Options

These options are set on the **Check Spelling** tab (**Tools>Options** menu).

Error display level

Note: This option must be set before you start recognition.

In the **Error display level** list, you may select the following values:

- **None** – recognition errors are not highlighted.
- **Standard** – unrecognized and uncertainly recognized characters are highlighted.
- **Thorough** – the same as **Standard**, words not found in the dictionary are highlighted too.

Note: The number of errors displayed in the **Text** window will change after you re-recognize the page.

Stop at words with uncertain characters

If you set this option, the spelling checker will stop at words with uncertain characters.

Stop at words not found in the dictionary

If you set this option, the spelling checker will stop at words not found in the dictionary. If a word is not found in the dictionary, it may have been recognized incorrectly.

Stop at compound words

If you set this option, the spelling checker will stop at those words not in dictionary which can be either made up according to the available morphology models or from the words in the dictionary.

Note: This option is set in the **Legacy Options** dialog. To open this dialog box, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.

Ignore words with digits and other non-alphabetic characters

If you set this option, the spelling checker will treat all words with digits and other characters not included in recognition language as correct, unless they contain uncertain characters.

Correct spaces before and after punctuation marks

If you set this option, the spelling checker won't stop if it finds incorrect spaces before or after punctuation marks; it will just correct the spaces automatically.

Note: This option is set in the **Legacy Options** dialog. To open this dialog box, click the **Legacy Options...** button on the **General** tab in the **Options** dialog.

Adding and Deleting Words to/from the User Dictionary

Adding words to the user dictionary

Enlarging the dictionary is a good way of increasing recognition quality. During recognition, ABBYY FineReader checks all words it comes across for possible dictionary entries. Therefore it makes sense to add new words that are likely to come up frequently (e.g. specialized terms, abbreviations, names etc.) to the user dictionary.

A distinctive feature of ABBYY FineReader's spelling checker is that a word is not only added to the dictionary in its original form, but its paradigm (i.e. the set of all its forms) is also added. This feature results in ABBYY FineReader being able to recognize a word in *all* its forms once it has been entered.

To add a word to the dictionary during the spelling check:

- Click the **Add** button in the **Check Spelling** dialog.

Set the following parameters in the **Primary Form** dialog:

1. **Part of speech** (Noun, Adjective, Verb, Uninflected).
2. If the word is to always begin with a capital letter, select the **Proper name** item. If you add an abbreviation, select the **Abbreviation** item.
3. The primary form of the word.

Click **OK**. The **Create Paradigm** dialog will open. ABBYY FineReader will ask you questions about the word forms in order to be able to construct the paradigm of the word you wish to add. Click **Yes** or **No** to answer these questions. If you make a mistake, click the **New** button to have ABBYY FineReader ask the question again. The constructed paradigm will be displayed in the **Paradigm** dialog.

Note:

1. If you do not wish paradigms to be created for the words you add, and want them to be entered uninflected instead, select the **Skip prompting for word forms** option (English dictionary only) on the **Check Spelling** tab (**Tools>Options** menu).
2. You may also add words when you view the list of added words. To do this, select the **View Dictionaries** item in the **Tools** menu. The **Select Dictionary** dialog will open. Select the language of your choice in the **Select Language** dialog and click **View**. The dictionary with the list of the added words will open. Add words by clicking on the **Add** button.
3. Paradigms can only be constructed for words added in the following languages: English, Italian, French, German (Old and New spelling), Russian, Spanish, and Ukrainian.

If the word you wish to add is already present in the dictionary, a notice to this effect will be issued. You may then wish to view its paradigm. If you think the existing paradigm is incorrect (this is often the case with homonymous words, for example), construct another one (click the **Add** button in the **Add Word** dialog).

Tip:

1. ABBYY FineReader allows you to import user dictionaries created by previous versions (5.0, 6.0 and 7.0).
2. ABBYY FineReader also allows you to import user dictionaries (*.dic) created using Microsoft Word 6.0, 7.0, 97, 2000 or later.

To import a dictionary:

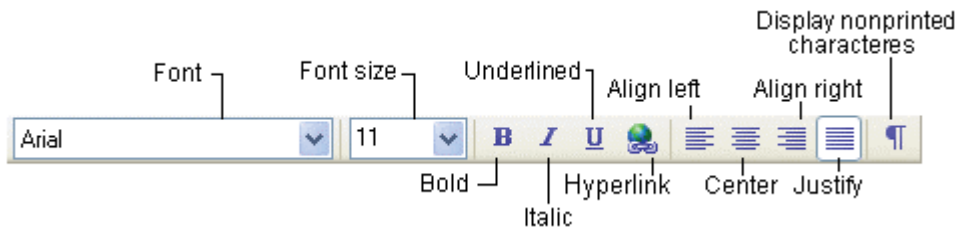
1. Select the **View Dictionaries** item in the **Tools** menu, then select the dictionary language, and click the **View** button.
2. Click the **Import** button in the opened dialog and select files with *.pmd, *.txt or *.dic extensions.

To delete a word from the dictionary:

1. Select the **View Dictionaries** item in the **Tools** menu. Select the language of your choice and click the **OK** button. A dialog will open.
2. Select the word you wish to delete and click the **Delete** button.

Editing Text in ABBYY FineReader

Note: If the ABBYY FineReader Text window does not display characters correctly (i.e. "?" or "?" can be seen in place of some or all of the letters), this means that your current font does not support your recognition language alphabet in full. Select a font that supports your entire recognition set (for example, Arial Unicode or Bitstream Cyberbit) on the **Save** tab (**Tools>Options** menu) in the **Fonts** group, and recognize the document again. See "Fonts for Recognition Languages that may be Displayed in Text Editor Incorrectly".





After a page is recognized, its text is displayed in the **Text** window. When you send your text to an external application, the text layout is retained according to the layout retention options chosen. Set these options on the **Save** tab (**Tools>Options** menu) and in the dialogs of the respective formats.

Uncertainly recognized characters are highlighted. To cancel this feature, unselect the **Highlight uncertain characters** item on the **View** tab (**Tools>Options** menu).

ABBYY FineReader editor features two document viewing modes: full mode (the full layout is displayed) and draft mode.

In full mode blocks with recognized text, tables and pictures are displayed exactly as they are to be found on the original image. The complete original layout, therefore, is retained: columns, tables, pictures, and dropped capitals (oversized letters that take up several lines of space in a paragraph). The block in which the pointer is currently located is the active block. If the pointer is moved using the arrow keys, the order of navigation between blocks is determined by their numbering on the original image. If the amount of text inside a particular block becomes too large for the block concerned (e.g. following editing), parts of other inactive blocks may become invisible. If this is the case, the borders of the block(s) concerned will be displayed with red markers. When a block is active, its borders are enlarged so as to display the entire block text.

The following text features are not displayed in draft mode: left indent; paragraph alignment (all paragraphs are aligned to the left); text and background color. A same-size font (12pt by default) is used throughout to display text in draft mode. Effects (bold, italic, underlined, superscript and subscript) are all retained.

Switch between draft and full modes by clicking the  (full mode) or  (draft mode) buttons in the **Text** window.

To change font size in draft mode:


1. Select the **Options** item in the **Tools** menu.
2. Set your preferred font size by selecting in the **Draft editor font size** item on the **View** tab.

The ABBYY FineReader built-in editor is supplied with the following text editing features:


Copy, cut, paste

1. Before you use the copy, cut, and paste commands, highlight the relevant text.
2. Follow the instructions below depending on the action you wish to carry out:

To copy the selection:

-  ● Either click the **Copy** button on the **Standard** toolbar, or
- Copy** button ● Select the **Copy** command in the **Edit** menu or local menu, or
- Press **CTRL+C**

To cut the selection:

-  ● Either click the **Cut** button on the **Standard** toolbar, or
- Cut** button ● Select the **Cut** command in the **Edit** menu or local menu, or
- Press **CTRL+X**

To paste the copied text:

-  ● Either click the **Paste** button on the **Standard** toolbar, or
- Paste** button ● Select the **Paste** command in the **Edit** menu or the local menu, or
- Press **CTRL+V**

Search and replace


To find a word or phrase in the text you are editing:

1. Perform one of the following actions:
 - Either select the **Find** item in the **Edit** menu, or
 - Press **CTRL +F**
2. The **Search** dialog will open. Type the word or phrase you wish to find in the **Find what** line of the dialog and set the search parameters.
Note: To search for the same word again using the same parameters, press **F3**.

To search and replace a word or phrase in the text you are editing:


1. Perform one of the following actions:
 - Either select the **Replace** item in the **Edit** menu, or
 - Press **CTRL+H**
2. The **Replace** dialog will open. Type the word or the phrase you want to find in the **Find what** line of the dialog, type the word or phrase that is to replace the search pattern in **Replace with** line, and set the search parameters.

Font effects

1. Click the word or highlight the text the font of which is to be changed.
2. Perform one of the following actions:
 - Either click the font-effect button (e.g. ) of your choice on the **Formatting** bar, or
 - Right-click the **Text** window and select **Character Properties** in the local menu. The **Character** dialog will open. Select the font type you wish to use and set the required font parameters in the dialog, or
 - Press **CTRL +B** – for boldface, **CTRL +I** – for italics, **CTRL +U** – to underline a word or text.

Note: You can also set the following additional text formatting parameters in the **Font** dialog: character spacing, character scale, and use of lowercase capitals. Keep in mind, however, that any formatting changes involving the latter will not be displayed in ABBYY FineReader's built-in text editor. These changes will only become visible once you export your document to an application that supports the latter formatting options (e.g. Microsoft Word).


Text alignment

1. Select the text you wish to align.
2. Perform one of the following actions:
 - Either click the alignment button (e.g. ) of your choice on the **Formatting** bar, or
 - Right-click the **Text** window and select the **Character Properties** item in the local menu. The **Character** dialog will open. Select the necessary item in the **Alignment** field.

Undo and redo

Perform one of the following actions:

To undo an action:

-  ● Either click the **Undo** button on the **Standard** toolbar, or
- Undo** button ● Select the **Undo** item in the **Edit** menu, or
- Press **CTRL+Z**.

To redo an undone action:



Redo button

- Either click the **Redo** button on the **Standard** toolbar, or
- Select the **Redo** item in the **Edit** menu, or
- Press **CTRL+Y**.

Editing Tables

The table editor provides you with tools to carry out the following:

- Merge cell or row contents
- Split cell contents
- Split row/column contents
- Delete cell contents

To merge cell or row contents:

- Hold down the **CTRL** button and select the cells or rows you wish to merge, followed by the **Merge Cells** or **Merge Rows** item in the **Image>Table Cells** menu.


To split cell contents:

- Select the **Split Cells** item in the **Image>Table Cells** menu.

Note: This command may only be applied to cells merged previously.

To split row or column contents:

- Select the  or  tool on the toolbar in the **Image** window, then click the row/column you wish to split or add a new horizontal/vertical separator to.

Tip: You can merge row contents by using the  tool or the **Merge Rows** command (menu **Image>Table Cells**).

To delete cell contents:

- Select the cell(s) you wish to delete in the **Text** window and press **DEL**.

Creating and Editing Hyperlinks

A hyperlink usually leads from one Web page to another. In addition, clicking a hyperlink may display an image, run a multimedia file, send an e-mail or launch an application.

FineReader 8.0 can recognize hyperlinks to:

- files, e.g. file://D:\MyDocuments\ABBYY FineReaderGuide.pdf;
- file-transfer protocol (FTP), e.g. ftp://;
- Web sites, e.g. http://www.abbyy.com;
- e-mail addresses, e.g. mailto:office@abbyy.com.

ABBYY FineReader will recognize the text of a hyperlink and recreate its destination address in the output document. Recognized hyperlinks are displayed in blue and underlined.


Note: You can change the color of a hyperlink on the **Save** tab in the **Options** dialog.

When viewing the recognized document in the **Text** window, rest the mouse cursor on a hyperlink to view its address. To follow a hyperlink, hold down the **Ctrl** button and click the hyperlink.


ABBYY FineReader editor allows you to:

- Change the text and address of a recognized hyperlink,
- Insert a hyperlink into a document,
- Delete a hyperlink from a document.

To change the text and address of a hyperlink:

1. Select the text of the hyperlink.
2. Click the  button on the toolbar of the **Text** window;
3. In the **Edit Hyperlink** dialog that opens, type a new text in the **Hyperlink text** field;
4. In the same dialog, specify the type of address in the **Link to** group:
 - Select **Web page** to link to an Internet page. In the **Address** field, specify the protocol and the URL of the page (e.g. *http://www.abbyy.com*);
 - Select **File** to link to a file. Selecting this option opens the **Open** dialog where you need to provide the name of the file to which the hyperlink will lead;
 - Select **E-mail** so that the user can send an e-mail message to the address in the hyperlink. In the **Address** field, specify the protocol and the e-mail address (e.g. *mailto:office@abbyy.com*).

To insert a hyperlink:

1. In the document, select the text you wish to use for the hyperlink.
If no text is selected, the address of the hyperlink will be displayed in the document;
2. Click the  button on the toolbar of the **Text** window;
3. In the **Edit Hyperlink** dialog that opens, type a new text in the **Hyperlink text** field;
4. In the same dialog, specify the type of address in the **Link to** group:

- Select **Web page** to link to an Internet page. In the **Address** field, specify the protocol and the URL of the page (e.g. *http://www.abbyy.com*);
- Select **File** to link to a file. Selecting this option opens the **Open** dialog where you need to provide the name of the file to which the hyperlink will lead;
- Select **E-mail** so that the user can send an e-mail message to the address in the hyperlink. In the **Address** field, specify the protocol and the e-mail address (e.g. *mailto:office@abbyy.com*).

To delete a hyperlink:

In the **Text** window, right-click the hyperlink you wish to delete and select **Delete Hyperlink** in the context menu.

Saving into External Applications and Formats

You can choose to save recognition results to a file, export them to an external application without saving them, copy them to the Clipboard or e-mail them in a supported file saving format. You can save specific pages or all of the pages in the document.

ABBYY FineReader can export recognition results to the following applications:

Microsoft Word 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) and 2003 (11.0); Microsoft Excel 6.0, 7.0, 97 (8.0), 2000 (9.0), 2002 (10.0) and 2003 (11.0); Microsoft PowerPoint 2002 (10.0) and 2003 (11.0); Corel WordPerfect 7.0, 8.0, 9.0, 10.0 (2002), 11.0 (2003) and 12.0; Lotus Word Pro 9.5, 97 and Millennium Edition; StarWriter 4.x, 5.x and 6.x, Adobe Acrobat/Reader (all versions).

General Information on Saving Recognized Text

You may:

- save recognized text using the **Save Wizard**,
- save opened or selected pages to a file or send them to an external application,
- save all the batch pages to a file or export them into an external application,
- save page image.



Click the **4-Save** button to send the recognition results to the application you need or save them to a file. The icon appearance changes to reflect the current saving mode. The **Save** button caption displays the name of the currently selected export application.

To save recognized text:

1. Click on the downward-pointing arrow to the right to the **4-Save** button and select the desired item in the local menu.

Note: To save a specific number of pages, select them prior to clicking the **4-Save** button.

After the export is finished, the **4-Save** button icon will change to reflect the previous action (sending the recognized text to an application, sending it by e-mail, copying it to the Clipboard or saving it to a file). The last export mode becomes the default for the **4-Save** button, so that clicking the icon will use that save option without going to the button's local menu.

Saving Options

The saving options are set before saving, on the the **Save** tab in **Options** dialog (**Tools>Options** menu). Some saving options can also be set in the **Save Wizard**, and **Save Pages**, **E-mail Pages** and **E-mail Images** dialogs. This section will outline:

- Fonts to use
- Save all or selected batch pages
- Recognized text saving modes

Fonts to use (when saving in RTF/DOC/Word XML, PPT or HTML format)

The fonts specified on the **Save** tab are used as the default when saving in RTF/DOC/Word XML, PPT or HTML formats. You can specify which fonts are used. To change fonts, go to the **Text** window or select other fonts on the **Save** tab in the **Fonts** group.

Save all or selected batch pages

You may choose to save all of the pages in a batch or only selected ones. To save specific pages, select them before saving.

Recognized text saving modes (when saving several batch pages at a time)

- **Create a separate file for each page** – The program saves each batch page as a separate file. The batch page number is automatically appended to the file name.
- **Name files as source images** – This option saves each page in a separate file and retains the name of the original image.

Note:

1. Pages that are un-related to the original image (e.g. scanned pages) will not be saved in this mode. A warning will be displayed when this type of page is encountered.
2. If consecutive batch pages share the same image as the original image or if all images have identical names, the program will treat the pages as a multi-page TIFF and save the text into a single file. If several pages have identical names but are not in consecutive order, the pages will be treated as individual image files, and the text will be saved in different files, with an index appended to their file names (_1, _2, etc.).

- **Create a new file at each blank page** – This option treats the entire batch as a set of page groups that contain a blank page at the end of each group. Pages from different groups are saved into different files with file names consisting of the user-specified name and index number: -1, -2, -3, etc.
- **Create a single file for all pages** – All (or all selected) batch pages are saved as a single file.

Saving the Recognized Text in RTF, DOC and Word XML Formats

Important! The option of saving in Word XML is only available for Microsoft Word 2003.

All saving options for **RTF**, **DOC** and **Word XML** formats are set on the **RTF/DOC/Word XML** tab in the **Formats Settings** dialog.

To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

Note: When saving text in RTF, DOC and Word XML formats, ABBYY FineReader uses the fonts set on the **Save** tab in the **Options** dialog (**Tools>Options** menu) or those you set during text editing in the **Text** window.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Page layout options
- Paper size options
- Text settings
- Picture settings

Retaining page layout

Layout retention modes are set in the **Retain layout** group. The following choices are available:

- **Columns, tables, paragraphs, fonts**
This option will retain the original layout in full, but in some cases there might be a slight difference from the original layout. Select this option if you are planning a lot of editing and re-formatting in the recognized text.
- **Original layout**
Select this option if you wish the recognition results to look exactly like the original document.
Note: This option will not allow a lot of editing in the recognized text. It is most suitable for short artistic or brochure-like documents.
- **Tables, paragraphs, fonts**
Select this option if you need the content of the original document, but do not need to retain the exact layout of the document.

Setting up the paper size

You may select the default paper size that will be used for saving in RTF, DOC or Word XML format. To do this, specify the required paper size in the **Default paper size** drop-down list.

Tips.

1. If you do not find a suitable paper size in the list, you can add your own – custom – paper size. In order to do this, select the **Add custom paper size** item from the list and in the dialog that appears specify the name, height and width for the custom paper size.
2. To ensure the recognition results fit the paper size, select the **Increase paper size if content does not fit** option. ABBYY FineReader will automatically select the most suitable paper size when saving the recognized text and pictures.

Text settings

Note that the default values of **Text settings** (an option is set or not) depend on the page layout retention mentioned above.

- **Keep line breaks**
This option saves the the original arrangement into lines to be retained the RTF/DOC/Word XML format.
- **Keep page breaks**
This option saves the original document page arrangement to be retained in RTF/DOC/Word XML format.
- **Retain text color**
This option saves the original character color to be retained.
Note: Word 6.0, 7.0 and 97 (8.0) have a limited text and background color palette. The original document colors may be replaced with the ones from the Word palette. Word 2000 (9.0) or later, on the contrary, retains the source document colors in full.
- **Remove optional hyphens**
This option removes the optional hyphen sign (–) from the recognized text. If the **Keep line breaks** option is set, the optional hyphen signs will be replaced with the hyphen signs (-).
- **Highlight uncertain characters**
Select this options if you wish to edit the recognized text in Microsoft Word rather than in the ABBYY FineReader **Text** window. If this option is set all uncertain characters will be highlighted in Microsoft Word window.
Tip. You may change the color of uncertain characters in the **View** tab of the **Options** dialog (**Tools>Options** menu).

- **Enable compatibility with Microsoft Word 95**
This option allows the recognition results to be saved in Microsoft Word 95.
Note: When saving in Microsoft Word 95, only the BMP image format is available for saving pictures.
- **Enable ABBYY FineReader's Zoom window in Microsoft Word 2003**
This option enables displaying ABBYY FineReader's **Zoom** window in Microsoft Word 2003. When saving results in Word XML, the recognized image can be viewed in the **Zoom** window integrated into Microsoft Word. This window presents the magnified image of the current line or portion of the document.

Picture Settings

If you wish to keep pictures in the recognized text, make sure that the **Keep pictures** option is set in the **Picture settings** group. If the recognized document contains many pictures, you can reduce the size of the resulting file: select the desired picture quality and format in the **Picture settings** group.

Quality

Three quality levels are available in the **Quality** drop-down list. Select:

- **High** if you are planning to print the recognition results.
- **Medium** if the recognition results are intended for viewing on the screen.
- **Low** if you are planning to place the recognition results on the Web.

The higher the value you choose from the **Quality** drop-down list, the higher will be the quality of the pictures you save. The size of the file is also affected by this value: the higher the value, the larger the file you get.

Tip. In order to tune the best 'size/quality' ratio, try to save the recognition results with different **Quality** values, and then open them in an image viewing application.

Format

As a rule, ABBYY FineReader selects the picture format automatically. To ensure that this is the case, make sure that the **(Automatic)** item is selected from the **Format** drop-down list.

If you wish to set up the format manually, select one of the following items:

- **JPEG, Color (for photos),**
This option is suitable for documents containing color scanned or digital photos.
- **JPEG, Gray (for photos),**
This option is suitable for scanned or digital photos saved in gray-scale mode.
- **PNG, Color (for charts, diagrams),**
This option allows you to save charts, diagrams or drawings while retaining their colors.
- **PNG, Gray (for charts, diagrams),**
This option is suitable for saving charts and diagrams in gray-scale mode.
- **PNG, Black and white.**
This option allows to you save pictures in black-and-white mode.

Saving the Recognized Text in XLS Format

All saving options for **XLS** format are set on the **XLS** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- **Ignore text outside tables**
This option allows you to save only tables and ignore the other recognition results.
- **Convert numeric values to numbers**
This option converts numbers into the "Numbers" format in the XLS file. Microsoft Excel may perform arithmetical operations with cells of this format.

Saving the Recognized Text in PDF Format

All saving options for **PDF** format are set on the **PDF** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Paper size options
- Save mode
- Tagged PDF option
- Picture and font settings
- Security options

Setting up the paper size

You may select the default paper size that will be used for saving in RTF, DOC or Word XML format. To do this, specify the required paper size in the **Default paper size** drop-down list.

Tips.

1. If you do not find a suitable paper size in the list, you can add your own – custom – paper size. In order to do this, select the **Add custom paper size** item from the list and in the dialog that appears specify the name, height and width for the custom paper size.
2. If you want to retain the size of the original page, select the **Keep original image size** option.

Save mode

ABBYY FineReader offers you four PDF creation modes:

- **Text and pictures only**
This option saves only the recognized text and the associated pictures. The page will be fully searchable and the PDF file size will be small.
- **Page image only**
This option saves the exact image of the page. This type of PDF will be virtually indistinguishable from the original but the file will not be searchable.
- **Text over the page image**
This option saves the background and pictures of the original document and places text over them. Usually, this PDF type requires more disk space than **Text and pictures only** and is fully searchable. In some cases there might be a slight difference from the original layout due to text being placed over the image.
- **Text under the page image**
This option saves the entire page image as a picture and places recognized text 'invisibly' underneath. Use this option to create a document with an absolutely perfect original layout and with full-text search capabilities.

Tagged PDF

In addition to contents, PDF files can contain information about the document structure such as logical parts, pictures, tables, etc. This structure is expressed via "PDF tags". A PDF file equipped with the tags may be reflowed to fit different screen sizes and will be displayed well on handheld devices.

If you wish to save recognized text to a tagged PDF file, select the **Enable Tagged PDF (compatible with Adobe Acrobat 5.0 or above)** option and ABBYY FineReader will automatically add PDF tags to the output PDF document.

Picture and fonts settings

If the recognized document contains many pictures, you can reduce the size of the resulting file: select the desired picture quality and format in the **Picture and font settings** group.

Quality

Three quality levels are available in the **Quality** drop-down list. Select:

- **High** if you are planning to print the recognition results.
- **Medium** if the recognition results are intended for viewing on the screen.
- **Low** if you are planning to place the recognition results on the Web.

The higher the value you choose from the **Quality** drop-down list, the higher will be the quality of the pictures you save. The size of the file is also affected by this value: the higher the value, the larger the file you get.

Tip. In order to tune the best 'size/quality' proportion, try to save the recognition results with different **Quality** values, and then open them in an image viewing application.

Format

As a rule, ABBYY FineReader selects the picture format automatically. To ensure that this is the case, make sure that the **(Automatic)** item is selected from the **Format** drop-down list.

If you wish to set up the format manually, select one of the following items:

- **JPEG, Color (for photos),**
This option is suitable for documents containing color scanned or digital photos.
- **JPEG, Gray (for photos),**
This option is suitable for scanned or digital photos saved in gray-scale mode.
- **LZW, Color (lossless),**
This option is suitable for graphics.
- **LZW, Gray (lossless),**
This option is suitable for graphics and gray images.
- **ZIP, Color (lossless),**
This option is suitable for pictures created with paint programs.
- **ZIP, Gray (lossless),**
This option allows you to save pictures in gray-scale mode.
- **CCITT4, Black and white.**
This option allows you to save pictures in black-and-white mode.

Font

When saving the recognized text to PDF format ABBYY FineReader uses the standard Adobe Acrobat fonts or system fonts. In the **Font** drop-down list, select:

- **Use standard fonts**
If this option is selected the PDF file refers to the standard Adobe Acrobat fonts: Times, Helvetica and CourierNew.

- **Use system fonts**

If this option is selected the PDF file refers to the standard fonts installed on the user's computer.

By default ABBYY FineReader embeds the fonts into the resulting PDF document. Embedded fonts ensure that the PDF document looks exactly like the original regardless of where it is viewed or printed. However, embedded fonts increase file size. If you do not need to embed fonts to your PDF documents, clear the **Embed fonts** option.

Tips.

1. When you save texts that use a non-Latin code page (such as Cyrillic, Greek, Czech, etc.), ABBYY FineReader uses the fonts provided by ParaType (www.paratype.com/shop).
2. Before you can edit PDF files that use non-Latin code pages in Adobe Acrobat, you must change the font of the current text section to a font installed on your computer.

Security

When saving recognized text to PDF format you can use passwords that will prevent the PDF document from being opened, printed or edited.

PDF Security Settings


When saving recognized text to PDF format you can use passwords that will prevent the PDF document from being opened, printed or edited.

The passwords and other restrictions are set in the **PDF Security Settings** dialog. To open this dialog, click the **PDF Security Settings** button on the **PDF** tab of the **Formats Settings** dialog.

You can specify a Document Open password or/and Permissions password.

Document Open password


A Document Open password prevents the users from opening the PDF document unless they type the password specified by the author. If you wish to add this password to your PDF document:

- Select the **Require password to open document** option, click  and in the the **Enter Document Open Password** dialog, type in the password and confirm it. The password you specified will be displayed with dots in the **Document Open password** field.

Permissions password

A Permissions password prevents the users from printing and editing the PDF document unless they type the password you specify. If some security settings are selected for the document the users will not be able to change these settings until they type the password you specify.

If you wish to add this password to your PDF document:

- Select the **Restrict printing and editing the document and its security settings** option, click  and in the the **Enter Permissions Password** dialog, type in the password and confirm it. The password you specified will be displayed with dots in the **Permissions password** field.

You can also enable or disable printing, editing or copying your PDF documents. These restrictions are defined in the **Permissions settings** group.

- The **Printing allowed** drop-down list enables/disables printing for the PDF document.
- The **Changes allowed** drop-down list specifies which editing actions are allowed in the PDF document.
- The **Enable copying text, images and other contents** option allows the users to select and copy text, pictures, etc. from your PDF document. If you wish to prevent the users from copying the content of the document, make sure that this option is cleared.
- The **Encryption level** drop-down list specifies the type of encryption for a password-protected document. The list allows you to select one of the three levels: the **Low (40 bit) – compatible with Acrobat 3.0 and above** item sets a low (40-bit RC4) encryption level; the **High (128 bit) – compatible with Acrobat 5.0 and above** item sets a high (128-bit RC4) encryption level, but Acrobat 3.0 users cannot open PDF documents with this encryption level; the **High (128 bit – AES) – compatible with Acrobat 7.0** item sets a high (128-bit RC4) encryption level, but Acrobat 6.0 (or earlier) users cannot open PDF documents with this encryption level.

Saving Recognized Text in HTML Format

All saving options for **HTML** format are set on the **HTML** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

Note: When saving text in HTML format, ABBYY FineReader uses the fonts set on the **Save** tab in the **Options** dialog (**Tools>Options** menu) or those you set during text editing in the **Text** window.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Page layout options
- Format options
- Text settings
- Picture settings
- Character encoding options

Retaining page layout

Layout retention modes are set in the **Retain layout** group. The following choices are available:

- **Original layout**
Select this option if you wish the recognition results to look exactly like the original document.
Note: This option will not allow a lot of editing in the recognized text. It is most suitable for short artistic or brochure-like documents.
- **Remove all formatting**
Only structure of tables and arrangement into paragraphs are retained. The text is set in the same font of the same size.
- **Retain font and font size**
Structure of tables, arrangement into paragraphs, font and font size are retained in this mode.

Format options

HTML formats available:

1. **Full (uses CSS and requires Internet Explorer 4.0 or later)** – the latest HTML format – HTML 4 – is used. HTML 4 supports all document layout retention types (the actual retention type used depends on the options set on the **Formatting** tab in the **Retain layout** group). The built-in style sheet is used.
Note: **Internet Explorer 4.0** or later is required for viewing a document saved in this mode.
2. **Simple (compatible with all (Internet-) browsers)** – HTML 3 format is used. The approximate document layout is retained i.e. the first line indent is not retained but the approximate font size is (HTML 3 format supports only a limited number of font sizes; ABBYY FineReader will choose the HTML 3 format font size that corresponds to the actual font size of your text). This HTML format is supported by all browsers (Netscape Navigator, Internet Explorer 3.0 and later).

Text settings

Note that the default values of **Text settings** (an option is set or not) depend on the page layout retention mentioned above.

- **Keep line breaks**
This option allows the the original arrangement into lines to be retained in HTML format.
- **Retain text color**
This option saves the original character color to be retained.
Note: Word 6.0, 7.0 and 97 (8.0) have a limited text and background color palette. The original document colors may be replaced with the ones from the Word palette. Word 2000 (9.0) or later, on the contrary, retains the source document colors in full.
- **Use solid line as page break**
This option saves the original arrangement into pages; the pages will be separated by a solid line.

Picture Settings

If you wish to keep pictures in the recognized text, make sure that the **Keep pictures** option is set in the **Picture settings** group.

If the recognized document contains many pictures, you can reduce the size of the resulting file: select the desired picture quality and format in the **Picture settings** group.

Quality

Three quality levels are available in the **Quality** drop-down list. Select:

- **High** if you are planning to print the recognition results.
- **Medium** if the recognition results are intended for viewing on the screen.
- **Low** if you are planning to place the recognition results on the Web.

The higher the value you choose from the **Quality** drop-down list, the higher will be the quality of the pictures you save. The size of the file is also affected by this value: the higher the value, the larger the file you get.

Tip. In order to tune the best 'size/quality' ratio, try to save the recognition results with different **Quality** values, and then open them in an image viewing application.

Format

As a rule, ABBYY FineReader selects the picture format automatically. To ensure that this is the case, make sure that the **(Automatic)** item is selected from the **Format** drop-down list.

If you wish to set up the format manually, select one of the following items:

- **JPEG, Color (for photos),**
This option is suitable for documents containing color scanned or digital photos.
- **JPEG, Gray (for photos),**
This option is suitable for scanned or digital photos saved in gray-scale mode.
- **PNG, Color (for charts, diagrams),**
This option allows you to save charts, diagrams or drawings with retaining their colors.
- **PNG, Gray (for charts, diagrams),**
This option is suitable for saving charts and diagrams in gray-scale mode.
- **PNG, Black and white.**
This option allows you save pictures in black-and-white mode.

Character encoding options

ABBYY FineReader detects the code page automatically. To change the code page, select the code page of your choice or the code page type in the **Character encoding** group.

Saving Recognized Text in PPT Format

When saving the recognition results to **PPT** format, ABBYY FineReader automatically retains full page layout.

All saving options for **PPT** format are set on the **PPT** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

Note: When saving text in PPT format, ABBYY FineReader uses the fonts set on the **Save** tab in the **Options** dialog (**Tools>Options** menu) or those you set during text editing in the **Text** window.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Text settings
- Picture settings

Text settings

- **Keep line breaks**
This option saves the original arrangement into lines in the PPT format.
- **Wrap text**
If line formatting is preserved, the recognized text will fit the width of the text block of the slide.

Picture Settings

If you wish to keep pictures in the recognized text, make sure that the **Keep pictures** option is set in the **Picture settings** group.

If the recognized document contains many pictures, you can reduce the size of the resulting file: select the desired picture quality and format in the **Picture settings** group.

Quality

Three quality levels are available in the **Quality** drop-down list. Select:

- **High** if you are planning to print the recognition results.
- **Medium** if the recognition results are intended for viewing on the screen.
- **Low** if you are planning to place the recognition results on the Web.

The higher the value you choose from the **Quality** drop-down list, the higher will be the quality of the pictures you save. The size of the file is also affected by this value: the higher the value, the larger the file you get.

Tip. In order to tune the best 'size/quality' ratio, try to save the recognition results with different **Quality** values, and then open them in an image viewing application.

Format

As a rule, ABBYY FineReader selects the picture format automatically. To ensure that this is the case, make sure that the **(Automatic)** item is selected from the **Format** drop-down list.

If you wish to set up the format manually, select one of the following items:

- **JPEG, Color (for photos),**
This option is suitable for documents containing color scanned or digital photos.
- **JPEG, Gray (for photos),**
This option is suitable for scanned or digital photos saved in gray-scale mode.
- **PNG, Color (for charts, diagrams),**
This option allows you to save charts, diagrams or drawings with retaining their colors.
- **PNG, Gray (for charts, diagrams),**
This option is suitable for saving charts and diagrams in gray-scale mode.
- **PNG, Black and white.**
This option allows you save pictures in black-and-white mode.

Important!

When saving results in the .PPT format, ABBYY FineReader creates special HTML files that contain the different parts of the presentation. To save the presentation as a single file, re-save it using PowerPoint (select **Save As** in the **File** menu and specify **PPT** as the saving format).

Saving the Recognized Text in TXT Format

All saving options for **TXT** format are set on the **TXT** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Text settings
- Character encoding options

Text settings

- **Keep line breaks**
This option saves the original arrangement into lines in the TXT format.
- **Append to end of existing file**
This option appends the text to the end of an already existing TXT file.
- **Insert page break character (#12) as page separator**
This option saves the original document page arrangement in TXT format.
- **Use blank line as paragraph separator**
If this option is selected the paragraphs will be separated by blank lines in the TXT file.

Character encoding options

ABBYY FineReader detects the code page automatically. To change the code page, select the code page of your choice or the code page type in the **Character encoding** group.

Saving the Recognized Text in DBF Format

All saving options for **DBF** format are set on the **DBF** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Text settings
- Character encoding options

Text settings

- **Append to end of existing file**
This option appends the text to the end of an already existing DBF file.

Character encoding options

ABBYY FineReader detects the code page automatically. To change the code page, select the code page of your choice or the code page type in the **Character encoding** group.

Saving the Recognized Text in CSV Format

All saving options for **CSV** format are set on the **CSV** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Text settings
- Character encoding options

Text settings

- **Ignore text outside tables**
This option allows you to save only tables and ignore the other recognition results.
- **Append to end of existing file**
This option appends the text to the end of an already existing TXT file.
- **Insert page break character (#12) as page separator**
This option saves the original document page arrangement in TXT format.
- **Field separator**
This field allows you to specify the character that will separate the fields in the CSV file.

Character encoding options

ABBYY FineReader detects the code page automatically. To change the code page, select the code page of your choice or the code page type in the **Character encoding** group.

Saving Recognized Text in LIT Format

All saving options for **LIT** format are set on the **LIT** tab in the **Formats Settings** dialog. To open this dialog, click the **Formats Settings** button on the **Save** tab of the **Options** dialog or press **CTRL+SHIFT+X**.

The following options enable you to customize the saving mode so that the resulting document is most suitable for later retrieval and processing:

- Text settings
- Picture settings

Text settings

- **Keep line breaks**
This option saves the original arrangement into lines in the LIT format.

- **Write tables as text**
This option converts table into text.
- **Retain text and background color**
This option allows you to retain the original color of the text and background.

Picture Settings

If you wish to keep pictures in the recognized text, make sure that the **Keep pictures** option is set in the **Picture settings** group. If the recognized document contains many pictures, you can reduce the size of the resulting file: select the desired picture quality and format in the **Picture settings** group.

Quality

Three quality levels are available in the **Quality** drop-down list. Select:

- **High** if you are planning to print the recognition results.
- **Medium** if the recognition results are intended for viewing on the screen.
- **Low** if you are planning to place the recognition results on the Web.

The higher the value you choose from the **Quality** drop-down list, the higher will be the quality of the pictures you save. The size of the file is also affected by this value: the higher the value, the larger the file you get.

Tip. In order to tune the best 'size/quality' ratio, try to save the recognition results with different **Quality** values, and then open them in an image viewing application.

Format

As a rule ABBYY FineReader select the picture format automatically. To ensure that this is the case, make sure that the **(Automatic)** item is selected from the **Format** drop-down list.

If you wish to set up the format manually, select one of the following items:

- **JPEG, Color (for photos),**
This option is suitable for documents containing color scanned or digital photos.
- **JPEG, Gray (for photos),**
This option is suitable for scanned or digital photos saved in gray-scale mode.
- **PNG, Color (for charts, diagrams),**
This option allows you to save charts, diagrams or drawings with retaining their colors.
- **PNG, Gray (for charts, diagrams),**
This option is suitable for saving charts and diagrams in gray-scale mode.
- **PNG, Black and white.**
This option allows you save pictures in black-and-white mode.

Saving the Page Image

1. Select a batch page.
2. Select File>Save Result>Save Images... The Save Image As dialog will open.
3. Select the disk or the folder where you want to save the file and the file format.
Note: You may want to save only some of the image areas enclosed by blocks (regardless of type). To do this, select the block or blocks you wish to save, and then check the **Save only selected blocks** checkbox in the **Save Image As** dialog. This is only an option when saving a single image. Next, enter the file name.
4. Click **OK**.

Tips:

- Saving several images to a single file (a multi-page TIFF)
 1. Select the images from the **Batch** window.
 2. Select the **Save Image As** item in the **File** menu. Select the TIFF format and the **Save as multi-page image file** option.
Note: If you save several page images from the **Batch** window as separate files (i.e. the images are not being saved as one multi-page TIFF), the file names will consist of the file name entered, the page number (4 digits), and the file suffix.
- Selecting a compression method for TIFF file format
TIFF is a universal and widely supported file format. Several compression methods formats are used with this format.
Compression reduces the size of an image. Compression methods are distinguished by the amount of compression and the quality of compressed images. ABBYY FineReader 8.0 supports the following compression methods:
- **ZIP** is a compression method suitable for images with large areas of single colors (e.g. screenshots), images created with paint programs, and black-and-white images. ZIP is a lossless method, i.e. it does not affect the quality of resulting images.
- **JPEG** is a compression method that is usually used for grayscale and color images, such as photographs. The JPEG is a lossy compression method which can greatly reduce the size of an image file. However, the image may appear blocky and acquire a mosaic look, and some shades of color may be lost.
- **CCITT** is a compression method most appropriate for black-and-white images which were scanned or created with a paint program. CCITT is a lossless compression method. CCITT Group 4 produces good compression for most types of images. CCITT Group 3 is used in fax machines.

- **Packbits** is a lossless compression method suitable for scanned black-and-white images.
 - **LZW** is a lossless compression method suitable for graphics and gray images.
- Note:** In ABBYY FineReader 8.0, this compression method is only available for saving pictures together with recognized text.

Adding Document Properties

Document properties contain the title of the document, the name of its author, its subject and keywords. You can use document properties to sort or find files. These properties are applied to all pages within the batch. To set or change document properties:

- In the **Batch** window, select the batch page whose properties you wish to set or change and then select **Document Properties** in the **File** menu,
- or
- In the **Batch** window, right-click the document whose properties you wish to set or change and then select **Properties** in the context menu.

In the dialog that opens, enter the title, the name of the author, the subject of the document and keywords.

Note: Not all file formats support document properties. The properties you specify will be added to documents in the following formats: DOC, XLS, HTM, PDF, and LIT.

Working with Batches

Batches are the main data depository in ABBYY FineReader, and contain scanned images, recognized text and other data. Most ABBYY FineReader settings are batch settings: scanning, recognition, saving options, etc. User patterns, user languages and user language groups are also property of a batch. New batches can be assigned the default batch settings, the settings of the current batch, or settings saved in an *.fbt file.

General Information on Working with Batches

ABBYY FineReader automatically creates a new batch upon starting. A batch may contain up to 9,999 pages.



Tip: Saving similar-type pages (e.g. pages from the same book, those written in the same language, or those with a similar layout) in the same batch is often useful, since it streamlines the work process.

The **Batch** window displays a list of the pages contained in the open batch. To view a page, click on its icon or double-click on its page number. All files related to this batch page will open in the appropriate windows, i.e. the text file in the **Text** window, the image file in the **Image** window, etc.

There are two ways of displaying pages in the Batch window:

Batch View	Description
Thumbnails	Batch pages are displayed as thumbnails (a miniature image of the original page). Additional icons appear on the thumbnails as the images are processed, to provide information about which actions have been performed on them (e.g. recognition, saving, etc.). Thumbnail images are particularly useful when searching for a particular batch page. To open an image, click on its thumbnail.
Details	This view provides detailed information about each batch page in the batch window and offers page lists organized by a user-specified feature. The batch window accommodates a large number of pages, which is useful when organizing large batches. Open a page by double clicking on it.

To choose the page view in the Batch window:

- Click either  or  on the **Standard** toolbar, or
- Right-click the **Batch** window and select the **Batch Window>...** item in the local menu.

To customize the **Batch** window to display specific features or sort according to certain criteria:


- Right-click the **Batch** window and select the **Batch Window>Customize** in the local menu. A dialog will open. Select the appropriate options on the **Thumbnails** and **Details** tabs of the dialog.

You may select several different pages, or a number of consecutive pages, or all of the batch pages in a row:

- **To select a number of consecutive pages**, hold down the **SHIFT** key and click the first and then the last page of the group you want to select.
- **To select several pages**, hold down the **CTRL** key and click the desired pages.
- **To select all batch pages**, activate the **Batch** window and choose the **Select All** item in the **Edit** menu or press **ALT+A**.

Creating a New Batch

To create a new batch:


- Select the **New Batch** item in the **File** menu, or
- Click on **New Batch** (.

Opening a Batch

ABBYY FineReader automatically creates a new batch at startup.

Note: To tell ABBYY FineReader to open the last open batch at startup, check **Open last batch at startup** on the **General** tab of the **Options** dialog (**Tools>Options**).

To open another batch:

1. Select the Open Batch in the File menu or click the Open Batch button (). The Open Batch dialog will open.
2. Select the appropriate folder in the **Open Batch** dialog.
When you open a batch, ABBYY FineReader automatically closes and saves the previous batch. Prior to exiting the program, save any new batch that might be needed in the future.

Batches can be opened directly from Windows Explorer:

- Right-click the batch folder (represented by the ) and select the **Open with ABBYY FineReader** item in the local menu. ABBYY FineReader will launch and open the chosen batch.

Adding Images to a Batch

- Select **Open PDF/Image** in the **File** menu or press **CTRL+O**.
- Select the desired image(s) in the **Open Image** dialog.
ABBYY FineReader will add the image to the open batch and copy the image to the batch folder.

Note: You can also add images directly from Windows Explorer:

1. Select an image file or group of files in Windows Explorer.
2. Right-click the selection and select **Open with ABBYY FineReader** from the local menu. If ABBYY FineReader is currently running, the selected files will be added to the current batch. Otherwise, ABBYY FineReader will be launched and the selected files will be added to a new batch.
This local menu item is only enabled in Windows Explorer if ABBYY FineReader 8.0 supports the file format.

Batch Page Number

All batch pages are numbered. A batch may contain up to 9,999 pages. The batch page numbers are displayed in the batch. You can renumber pages directly in the **Batch** window or from the **Renumber Pages** dialog.

To renumber pages directly in the Batch window:

1. Click a page in the **Batch** window or press **F2**.
2. Enter the new page number.

Once the page number has been changed, all pages in the **Batch** window will be re-ordered to reflect the new numbering.

Note: Double-clicking a page number opens the page.

To renumber pages in the Renumber Pages dialog:

1. Select a single page or several pages.
2. Select **Renumber Pages** in the **Batch** menu.
3. Set the new number for the first page selected (the page with the lowest number).

Note:

1. To renumber all batch pages, select **All Pages** in the **Renumber Pages** dialog.
2. To renumber only some portion of a batch:
 - Select the pages you wish to renumber in the **Batch** window.
 - Choose **Selected pages** from the Renumber Pages dialog.
3. To renumber selected pages continuously, select the **Continuous page numbering** option. For example: The renumbering option would cause pages numbered 2, 5, and 6 (assuming 1 was chosen as the first number) to be renumbered as 1, 2, and 3. Otherwise (i.e. if the **Continuous page numbering** option is not selected), on renumbering page numbers 2, 5, and 6 would become 1, 5, 6. The first page has been assigned the chosen number, but the remaining pages have retained their original numbers.

Note: If you renumber only certain batch pages, and allocate a number to a page that has been used, a warning will be issued and the operation will be cancelled.

Saving a Batch

To save a batch:

- Select **Save Batch As** in the **File** menu.
- In the **Save Batch As** dialog, specify the name of the batch and the desired storage location.

Closing a batch page or the whole batch

To close a batch page:

- Select Close Current Page in the Batch menu.

To close a batch:

- Select **Close Batch** in the **File** menu.

Deleting a Batch

Note: When a batch is deleted, all of its contents (including image and text pages, related files, user patterns, user languages, etc.) will be deleted, leaving the folder empty.

To delete a batch:

- Select **Delete Batch** in the **Batch** menu.

To delete a batch page:

1. Select the page(s) you wish to delete in the **Batch** window.
2. Select **Delete Page from Batch** in the **Batch** menu or press **DEL**.

Batch Settings

To save batch settings in a file:

- Click on **Save Options...** on the General tab (Tools>Options). The Save Options As dialog will open.
- Enter the file name.

The following settings will be saved: the **Scan/Open**, **Read**, **Check Spelling** and **Save** tab settings, and the settings specified in the **Formats Settings** dialog box. The user languages, user language groups and user patterns will also be saved in the file. To apply the options set to all new batches, check **Apply this options set to new batches** in the **Save Options As** dialog.

To return to the default settings:

- Click on Reset to Defaults on the General tab.

To load the settings:

- Click **Load Options...** on the **General** tab and select the ABBYY FineReader option set (*.fbt) file that contains the desired settings.

Full-Text Search in Recognized Batch Pages

Important! You need Internet Explorer 4.0 or later to access this option.

ABBYY FineReader allows you to search all recognized pages for words in every possible grammatical form. The search pattern may consist of one or several words. The search term may be in any form (for languages with dictionary support), and the search process will identify the indicated words anywhere within the text (no matter how far apart) and in any order.

To do a full-text search:

1. Select **Advanced search** in the **Edit** menu or press **ALT+F3**.
2. The **Search** window will open below the **Zoom** window.
3. Enter the desired text in the **Find what** field. You can also paste the Clipboard contents into the field or select a previous search from the drop-down list.
4. Click **Find**.

The **Search results** window will display the list of batch page numbers that contain all of the words from the **Find what** field. The date that each identified page was modified will be displayed and the first page section to contain the search pattern will be highlighted. Click on a page number to open it in the **Image**, **Text** and **Zoom** windows. The found words will be highlighted in color in all three windows.

Note: The search function cannot locate specialized characters, such as end-of-line characters or paragraph marks.

Automated Tasks

Very often an OCR process involves a number routine tasks such as scanning, recognition, and saving the results in a particular format. ABBYY FineReader 8.0 offers tools for automating routine tasks for similar documents.

An **automated task** is a sequence of steps, each corresponding to a particular processing routine. Automated tasks are launched from the menu of the **Scan&Read** button.

ABBYY FineReader 8.0 already includes three automated tasks which are ready for use and do not require any additional settings. You can also use an **Automation Wizard** to create your own automated tasks to suit your needs.

Running an Automated Task

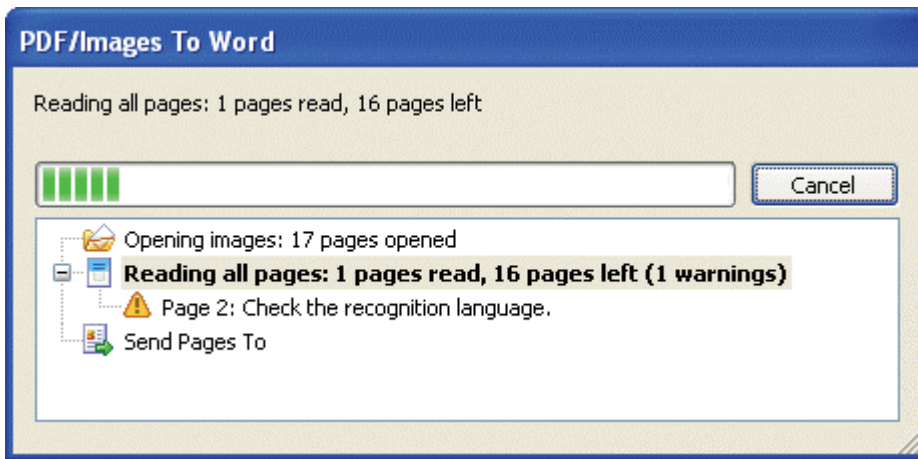
To run an automated task:

- Click the arrow at the right of the **Scan&Read** button and select the required automated task, or
- Press **Ctrl+W** (for a previously selected task), or
- Select the required task from the **Process>Automated Tasks** menu, or
- Open the **Automation Manager** dialog, select the required automated task and click **Run**.

The selected task will be launched and executed. ABBYY FineReader will use the options you selected in the **Options** dialog (select **Tools>Options** to open the dialog).

Note: If an automated task is programmed to add images to a batch which already contains some other images, ABBYY FineReader will process only the newly added images. If an automated task is programmed to process an existing batch, ABBYY FineReader will process all the images in this batch.

While the automated task is running, a progress dialog is displayed which lists all the steps in the task and provides tips and warnings for each step (if any).



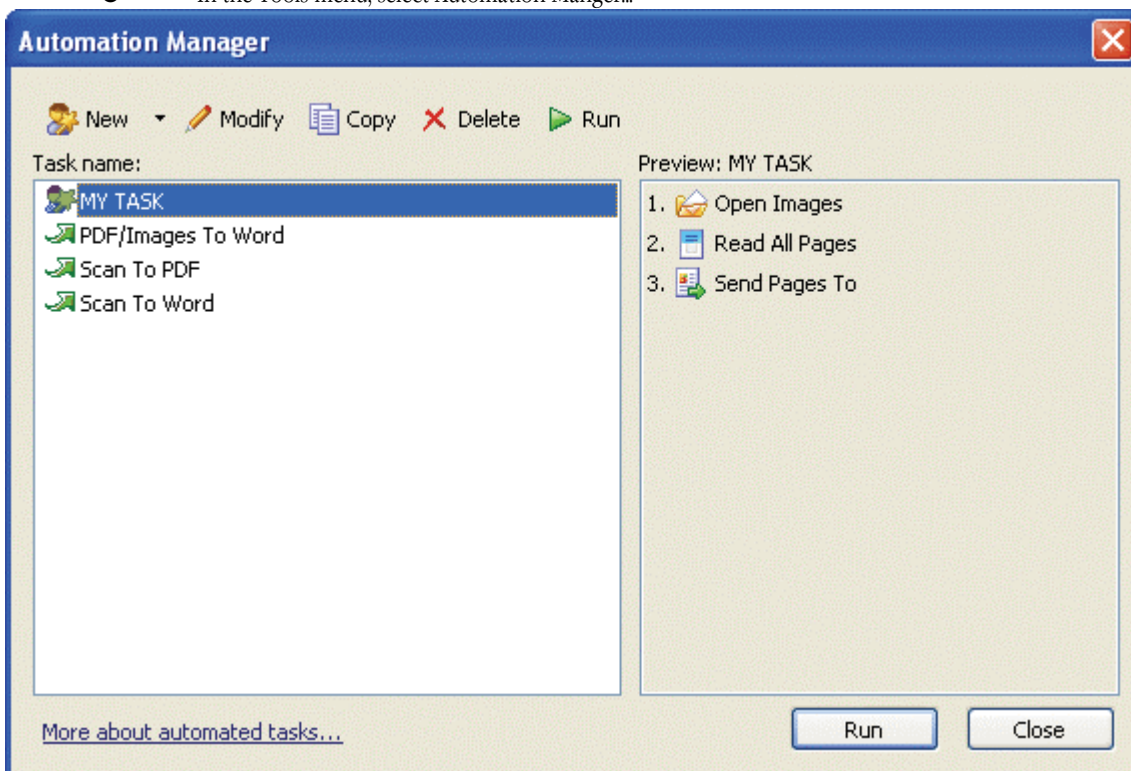
Note: If you want an automated task to use options which you do not normally use when recognizing documents, you can create a set of **custom batch settings** and load these settings before running the automated task. To create a set of custom batch settings, make the necessary settings in the **Options** dialog and on the **General** tab of this dialog, click **Save Options...** Next time before you run an automated task you can load the saved option set by click the **Load Options...** button.

Managing Automated Tasks

The **Automation Manager** allows you to run an automated task, to create and modify automated tasks, and delete custom automated tasks which you no longer use.

To run the Automation Manager:






- Click the arrow at the right of the **Scan&Read** button and select **Automation Manager...**, or
- Press **Ctrl+T**, or
- In the Process menu, select Automated Tasks>Automation Manger..., or
- In the Tools menu, select Automation Manger...



Use the buttons on the **Automation Manager** toolbar to create, modify, delete or run automated tasks.

The left-hand pane lists the available automated tasks. The automated tasks shipped with ABBYY FineReader are marked with and custom automated tasks are marked with . Automated tasks which cannot be run on your computer are marked with . Clicking an automated task in the left-hand pane displays its steps in the right-hand pane.

Note: If you want recognized text to be sent to another application, this application must be installed on your computer. Automated tasks which are programmed to send recognized text to applications which are not installed will not run. Such automated tasks will not be displayed in the drop-down list at the right of the **Scan&Read** button or in **Process>Automated Tasks**.


Button name	Button description
 New	Creates a new automated task. The Automation Wizard will help you select the required steps and make the settings.
Export...	Exports an automated task to a file which can be used on other computers. In the Export Automated Task dialog that opens, specify an *.fta file to which the automated task will be saved. Note: By default, ABBYY FineReader saves automated tasks in %UserProfile%\Local Settings\Application Data\ABBYY\FineReader\8.00 .
Import...	Imports an automated task. In the Import Automated Task dialog, select an *.fta file you wish to import and a folder where it will be stored.
 Modify	Modifies the properties of the selected automated task. The Automation Wizard will help you select the required properties. Note: You cannot modify automated tasks shipped with ABBYY FineReader. However, you can copy such tasks and then modify these copies.
 Copy	Copies an automated task. The copy of an automated task will be added to the list of available automated tasks immediately after its original.
 Delete	Deletes an automated task. Note: You cannot delete automated tasks shipped with ABBYY FineReader.
 Run	Runs an automated task.

Tip: To rename an automated task, right-click the task and select **Rename** in the local menu.

Automated Tasks Shipped with ABBYY FineReader 8.0

ABBYY FineReader ships with three automated tasks which cover all the steps required to get an image of a document, recognize it and save the recognized text. These automated tasks do not include the step of checking the recognition results. Your copy ABBYY FineReader already includes the following three automated tasks:

PDF/Images To Word

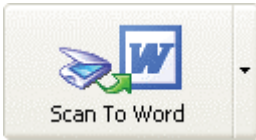


This automated task:

1. Opens image files (JPEG, TIFF, etc.) or PDF files. When you run the **PDF/Images To Word** task, the **Open Image** dialog will open. Select the files to be processed and click **OK**. The selected images will be added to the current ABBYY FineReader batch.
2. Recognizes the images.
3. Sends the recognized text to Microsoft Word.
A new Microsoft Word document will be opened and the recognized text will be placed in this document.

Note: This automated task requires Microsoft Word to be installed on your computer.

Scan To Word




This automated task:

1. Scans paper documents.
The images of the scanned documents will be added to the current ABBYY FineReader batch.
2. Recognizes the images.
3. Sends the recognized text to Microsoft Word.
A new Microsoft Word document will be opened and the recognized text will be placed in this document.

Note: This automated task requires Microsoft Word to be installed on your computer.

Scan to PDF



This automated task:

1. Scans paper documents.
The images of the scanned documents will be added to the current ABBYY FineReader batch.
2. Recognizes the images.
3. Sends the recognized text to Adobe Acrobat/Reader.
A new PDF document will be opened and the recognized text will be placed in this document.

Custom Automated Tasks

You can create your own automated tasks to suit your document processing needs. You can add more steps to your automated tasks than are available in the five automated tasks which ship with ABBYY FineReader. For example, you can add the step of checking the recognition results or add multiple saving steps to save your document in several formats.

The sequence of steps in an automated task must follow the usual order of processing a document in ABBYY FineReader: *acquiring an image – recognition – checking the results – saving the results*.

Each new step is added with default properties. You can change the properties to suit your needs.

An automated task may contain main and additional steps.

Main steps

The main steps are: acquiring images, recognition, and saving. One automated task may include only one step of acquiring images, one recognition step, and several saving steps.

● **Acquiring images**

This is always the first step in an automated task. At this step, ABBYY FineReader gets images to be processed.

Step	Property	Description
Scan Images	ABBYY FineReader uses the current batch settings to scan the images.	Scans paper documents.
Open Images	Prompt for image file names when task starts (default)	When you run the task, ABBYY FineReader will prompt you to select image files and add them to the current batch. In the Open Images dialog, select the files to be processed and click OK .
	Process images from folder	When you run the task, ABBYY FineReader will open the folder you specified in the field below and add all the images from this folder to the current batch. Select the Include all subfolders box if you wish ABBYY FineReader to look for images in all the subfolders.
Open Batch	Prompt for batch name when task starts (default)	When you run the task, ABBYY FineReader will prompt you for a batch name. In the Open dialog that opens, select the batch to be processed.
	Use current batch	When you run the task, ABBYY FineReader will start processing the images in the current batch.
	Use this batch	When you run this task, ABBYY FineReader will start processing the images from the batch you specified in the field below.

● **Layout analysis**

Step	Property	Description
Load Blocks Template	Prompt for block template when task starts (default)	When you run the task, ABBYY FineReader will prompt you for a block template. Browse to the required template file in the Open dialog and click OK .
	Use this block template	Provide the path to the template file to be used.
	Check and adjust blocks manually	Once the program has analyzed the layout and drawn the necessary blocks, you can review and adjust them manually.
Analyze Layout	Analyze layout automatically then adjust blocks manually (default)	Once ABBYY FineReader has acquired the images, it will analyze them and draw the necessary blocks. Then you can review and adjust them manually.
	Draw blocks manually	Once ABBYY FineReader has acquired the images it will ask you to draw the necessary blocks manually.

● **Recognition**

At this step, ABBYY FineReader recognizes the images.

Step	Property	Description
Read All Pages	No properties	Automatically recognizes the images in the specified batch or folder.

● **Checking the recognition results**

Step	Property	Description
Check Results	Check spelling	Once ABBYY FineReader has recognized the text, the Check Spelling dialog will open.
	Review results without spell check	The recognized pages will be displayed in the Text window where you can review it without running the spell-checker.

● **Saving**

At this step, ABBYY FineReader saves the text to a file or sends it to an application of your choice. An automated task may include several saving steps.

Step	Property	Description
Save Pages	Prompt for output file names when saving (default)	ABBYY FineReader will open the Save Pages dialog prompting you to select the file and saving options.

	Save with specified names and to specified location	<p>If you select this property, you need to specify the following:</p> <ol style="list-style-type: none"> 1. Output folder Specify the folder where the file(s) containing the recognized text will be saved. Check the Create time-stamped subfolder box if you wish ABBYY FineReader to create a new subfolder each time you run this task. This option is useful if you do not wish to specify the folder manually each time you run the task. 2. Save as type From the drop-down list, select file format. 3. File options: <ul style="list-style-type: none"> ● Create a single file for all pages – Saves all the pages (or all the selected pages) of the batch to one file. ● Create a separate file for each page – Saves each page to a separate file. ● Create a new file at each blank page – ABBYY FineReader uses blank pages to divide the pages into groups. For each group, a separate file is created into which all the pages of the group are saved. ABBYY FineReader will name the created files by adding -1, -2, -3, etc. to the file name specified in the Name field. ● Name files as source images – Saves each page to a separate file named as the original image. 4. File name.
Save Images	Prompt for image file names when saving (default)	ABBYY FineReader will open the Save Image As dialog prompting you to select the file and saving options.
	Save images with specified names and to specified location	<p>If you select this property, you need to specify the following:</p> <ol style="list-style-type: none"> 1. Output folder Specify the folder where the file(s) containing the images will be saved. 2. Save as type From the drop-down list, select file format. Select the Save as one multi-page image file option if you wish to save all the images into one multi-page file. Note: This option is available only for TIFF and PDF file formats. 3. File name.

Additional steps

Additional steps of an automated task are used to send the recognized text to an external application, attach the acquired image or the recognized text to an e-mail message, and copy ABBYY FineReader batches.

● **Sending pages to another application**

Step	Property	Description
Send Pages To	Save Wizard (default)	Use the Save Wizard or select the desired application from the drop-down list. The recognized text will be placed into a new file and opened in the application of your choice.

● **Sending the image or recognized text as an e-mail attachment**

Step	Property	Description
E-mail Pages	Attach as type	Select the required file format from the drop-down list. The recognized text will be saved to a file of the selected format. See full list of the image file formats supported by ABBYY FineReader see in "Supported Document Saving Formats". Note: The recognized pages can be saved into several files according to your choice made in File options .
	File options	From the drop-down list, select one of the options. The following choices are available: <ul style="list-style-type: none"> ● Create a single file for all pages All pages are saved to one file. <i>The option is set by default.</i> ● Create a separate file for each page Each page is saved to a separate attached file. For each page, a separate file is created into which all the pages are saved. ABBYY FineReader will name the created files by adding -0001, -0002, -0003, etc. to the default file name. ● Create a new file at each blank page ABBYY FineReader uses blank pages to divide the pages into groups. For each group, a separate file is created into which all the pages of the group are saved. ABBYY FineReader will name the created files by adding -1, -2, -3, etc. to the default file name. ● Name files as source images Each page is saved as a separated attached file named as the original image.

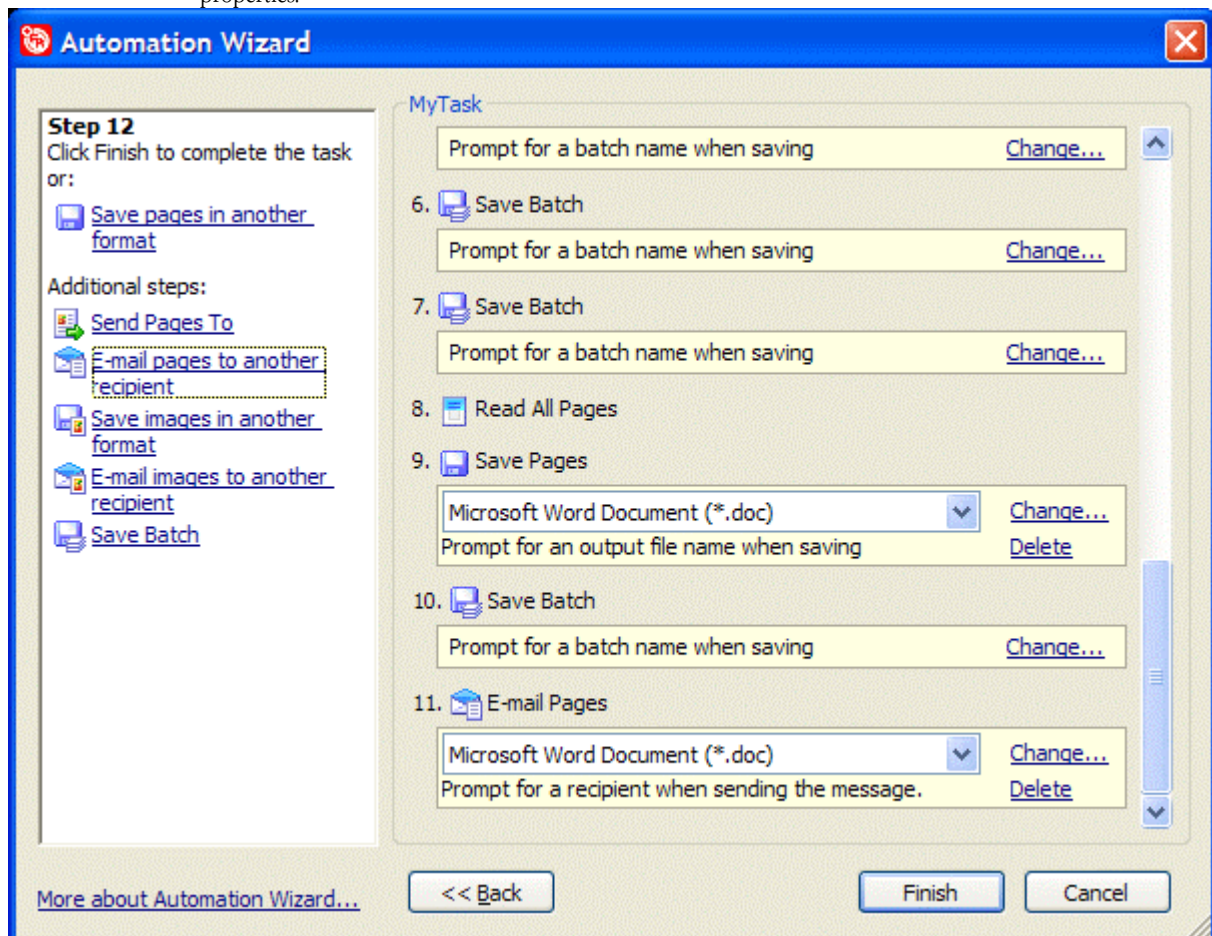
E-mail Images	Attach as type	Select the desired file format from the drop-down list. The selected images will be attached to an e-mail message. See the full list of the image file formats supported by ABBYY FineReader in "Supported Image Formats".
	Send as one multi-page image file	Select this option if you wish to save all the images into one multi-page file. Note: This option is available only for TIFF and PDF file formats.
	Name	Specify the file name. Note. If you save images to separate files (the Save as one multi-page image file option is not selected), ABBYY FineReader will add the page number or page group number (0001, 0002, etc.) to the name of each file.

● **Saving the batch**

Step	Property	Description
Save Batch	Prompt for batch name when saving (default)	At this step, the Save Batch As... dialog will open where you have to specify a folder where the batch will be stored.
	Save batch to	Browse to the folder where the batch will be stored.

Automating a Task

1. Start the Automation Manager:
 - Select the **Automation Manager** command from the drop-down list at the right of the **Scan&Read** button, or
 - Press **Ctrl+T**, or
 - Select **Automated Tasks>Automation Manager** from the **Process** menu, or
 - Select the **Automation Manager** command from the **Tools** menu.
2. In the **Automation Manager** dialog, click **New**.
3. In the dialog that opens, enter a name for the new automated task.
4. The **Automation Wizard** will open. The wizard will guide you through the automation steps and their properties.



The left-hand pane of the **Automation Wizard** displays the list of available steps. As you select steps in this list, new steps may become available or, conversely, some steps may become unavailable. The right-hand panel displays the selected steps and their properties.

5. Select a step in the left-hand pane. The selected step will be displayed in the right-hand panel.
6. The property of a step is displayed in a yellow field below. If you wish to change the default property, click the **Change...** link to the left and select a new property.
7. The saving steps have the **Delete** link that allows you to remove an unwanted step from your automated task.
Note: The scanning/opening, recognition and page layout analysis steps cannot be removed independently. To remove these steps from the automated task, use the **Back** button.
8. Once you have added all the necessary steps to your automated task and selected their properties, click **Finish**.

The new task will be added to the list of available tasks in the **Automation Manager** and to the drop-down list of tasks at the right of the **Scan&Read** button.

Chapter 2

ABBYY Screenshot Reader

ABBYY Screenshot Reader is an easy-to-use application which allows you to create screenshots and recognize texts.

ABBYY Screenshot Reader features:

- OCR of text in any section of the computer screen.
- OCR of tables in any section of the screen.
- Creating screenshots of any section of the screen.
- Saving OCR results to a file, copying them to the Clipboard or sending them to another application.

ABBYY Screenshot Reader has a straightforward and intuitive interface, which means that you do not need any specialist knowledge to be able to make screenshots and recognize text in them. Simply open any window of any application and select the section of the computer screen which you wish to "photograph".

Note: ABBYY Screenshot Reader is available to all users of ABBYY FineReader 8.0 Corporate Edition and to registered users of ABBYY FineReader 8.0 Professional Edition.

Chapter Contents

- Installing and starting ABBYY Screenshot Reader
- ABBYY Screenshot Reader toolbar
- Capturing text and tables from the computer screen
- Making screenshots
- Additional options

Installing and Starting ABBYY Screenshot Reader

Installing ABBYY Screenshot Reader

By default, ABBYY Screenshot Reader is installed on your computer together with ABBYY FineReader 8.0. If during custom installation you chose not to install ABBYY Screenshot Reader, you can install the application by following the instructions below:

1. On the Windows taskbar, click **Start** and then select **Settings>Control Panel**.
2. In the list of installed programs, select **ABBYY FineReader 8.0** and then click **Change**.
3. In the Custom Setup dialog box, select ABBYY Screenshot Reader.
4. Follow the instructions of the setup program.

Activating ABBYY Screenshot Reader

Note: Activation is only required for users of ABBYY FineReader 8.0 Professional Edition.

You must be a registered ABBYY FineReader user to be able to use ABBYY Screenshot Reader. Upon registration, you will be provided with an **ABBYY Screenshot Reader Code**. To activate ABBYY Screenshot Reader, enter this code into the dialog that opens when you launch the application and click **OK**.

Starting ABBYY Screenshot Reader

To start ABBYY Screenshot Reader, go to Start>Programs>ABBYY FineReader 8.0 and select ABBYY Screenshot Reader.

ABBYY Screenshot Reader Toolbar

When you start ABBYY Screenshot Reader, its toolbar is displayed.



The **ABBYY Screenshot Reader** toolbar contains tools for recognizing text and tables on the screen of your computer, creating screenshots of selected areas on the screen, and for setting up ABBYY Screenshot Reader.



Clicking this button turns on the selection tool which allows you to select an area on the screen. The selected area will be enclosed in a frame and depending on your settings the program will either automatically start recognizing the text in the selected area or create a screenshot of the selected area. The resulting text or screenshot will be saved to a file, copied to the Clipboard or sent to another application, depending on which option you choose from the drop-down list to the right.



In this drop-down list, select a screen object to capture and a destination where it should be saved.



Opens the **Options – ABBYY Screenshot Reader** dialog where you can select a recognition language, toggle between the two display modes of the **ABBYY Screenshot Reader** toolbar, and select a sound and/or a message which ABBYY Screenshot Reader will use to signal that a screenshot has been copied to the Clipboard.



This button toggles between two display modes of the **ABBYY Screenshot Reader** toolbar. If you select  , the **ABBYY Screenshot Reader** toolbar will always be displayed above the windows of the other running applications.

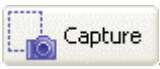

Capturing Text and Tables from the Computer Screen

ABBYY Screenshot Reader can recognize text and tables on the screen of your computer and save the results to a file, copy them to the Clipboard or send them to another application.

To recognize text:

1. In the drop-down list on the ABBYY Screenshot Reader toolbar, select one of the following:
 Text to Clipboard
 Table to Clipboard
 Text to Microsoft Word

Table to Microsoft Excel
Text to File
Table to File


2. Click . The mouse cursor will change to .
3. Place the mouse cursor in a corner of the area you wish to select. Hold down the left mouse button and drag the cursor diagonally to the opposite corner of the area you wish to select. The selected area will be enclosed in a frame and the program will automatically start the OCR procedure.
4. Once the text has been recognized, it will be saved to a file, copied to the Clipboard or sent to another application, depending on the action you selected in the drop-down list on the ABBYY Screenshot Reader toolbar. If you selected **Text to File** or **Table to File**, the **Save As** dialog will open. In this dialog, provide the name and format of the file into which the text must be saved.

Note: To cancel the recognition, click **Stop** or press **Esc**.

What do I do if ...

I work with texts written in several languages

Before starting the recognition procedure, make sure that the language you selected in **Options – ABBYY Screenshot Reader** dialog is the same as the language of your text. Select a different recognition language if required.

Note: To open the **Options – ABBYY Screenshot Reader** dialog, click . **The text on the screen appears to be written in several languages**

In the **Options – ABBYY Screenshot Reader** dialog, select the **(Select multiple languages...)** item in the Recognition language drop-down list.

Important! Using multiple languages can lower recognition quality. We do not recommend using more than two or three languages.

I have not found the required language in the list

In the **Options – ABBYY Screenshot Reader** dialog, select the **(Choose more languages...)** item in the Recognition language drop-down list.

If you cannot find the required language in the list of available languages:

- This language may be unsupported by ABBYY FineReader 8.0. See "Languages Supported by ABBYY FineReader" for the full list of recognition languages.
- This language may have been disabled during custom installation:
 1. On the Windows taskbar, click Start and then select Settings>Control Panel.
 2. In the list of installed programs, select ABBYY FineReader 8.0 and then click Change.
 3. In the Custom Setup dialog box, select the required language.
 4. Follow the instructions of the setup program.

The selected area contains both text and pictures

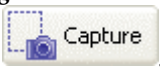

On the **ABBYY Screenshot Reader** toolbar, select the **Image to ABBYY FineReader** item. The image of the selected area will be added to a ABBYY FineReader batch. If ABBYY FineReader is running on your computer, the image will be added to the current batch. Otherwise, a new batch will be created and the image will be added to this new batch.

Making Screenshots

ABBYY Screenshot Reader can create screenshots of selected areas on the screen of your computer and save them to a file, copy them to the Clipboard or send them to ABBYY FineReader.

To create a screenshot:

1. In the drop-down list on the **ABBYY Screenshot Reader** toolbar, select one of the following:
Image to Clipboard
Image to ABBYY FineReader (select this option if the screen area contains both text and pictures)
Image to File

2. Click . The mouse cursor will change to .
3. Place the mouse cursor in a corner of the area you wish to select. Hold down the left mouse button and drag the cursor diagonally to the opposite corner of the area you wish to select. The selected area will be enclosed in a frame.
4. The screenshot will be saved to a file, copied to the Clipboard or added to a ABBYY FineReader batch, depending on the action you selected in the drop-down list on the ABBYY Screenshot Reader toolbar. If you selected **Text to File**, the **Save As** dialog will open. In this dialog, provide the name and format of the file into which the text must be saved.

Note: To cancel screenshot creation, press **Esc**.

Additional Options

You can select additional options in the **Options – ABBYY Screenshot Reader** dialog. To open the **Options – ABBYY Screenshot**

Reader dialog, click  on the **ABBYY Screenshot Reader** toolbar. In this dialog you can:

- Select the recognition language to match the language of the text in the selected screen area.
- Select the **Always on top** box, to make the ABBYY Screenshot Reader toolbar appear always above the windows of other running applications.
- Select **Play sound if data has been copied successfully** to make ABBYY Screenshot Reader play a sound after the data has been copied to the Clipboard.
- Select **Show message if data has been copied successfully** to make ABBYY Screenshot Reader display a notification message after the data has been copied to the Clipboard.

Chapter 3

ABBYY Hot Folder & Scheduling

ABBYY FineReader 8.0 now includes **ABBYY Hot Folder & Scheduling**, a scheduling agent. ABBYY Hot Folder & Scheduling allows you to select a folder with images and set the time for processing images in this folder. For example, you can schedule your computer to recognize images overnight.

To set up a "hot" folder, you need to select image opening, recognition, and saving options, specify how often ABBYY FineReader should check the folder for new images (at regular intervals or only once), and set the start time.

Chapter Contents

- Installing and starting ABBYY Hot Folder & Scheduling
- Hot ABBYY Folder & Scheduling main window
- Setting up a hot folder
- Hot folder log file
- Additional options

Installing and Running ABBYY Hot Folder & Scheduling

By default, ABBYY Hot Folder & Scheduling is installed on your computer together with ABBYY FineReader 8.0. If during custom installation you chose not to install Hot Folder & Scheduling, you can install the application by following the instructions below:

1. On the Windows taskbar, click **Start** and then select **Settings>Control Panel**.
2. Double-click the **Add or Remove Programs** icon in the **Control Panel** window.
3. In the list of installed programs, select **ABBYY FineReader 8.0** and then click **Change**.
4. In the Custom Setup dialog box, select ABBYY Hot Folder & Scheduling.
5. Follow the instructions of the setup program.

Running ABBYY Hot Folder & Scheduling

To run ABBYY Hot Folder & Scheduling:

- Go to Start>Programs>ABBYY FineReader 8.0 and select ABBYY Hot Folder&Scheduling

or

- In ABBYY FineReader 8.0, select ABBYY Hot Folder & Scheduling from the Tools menu

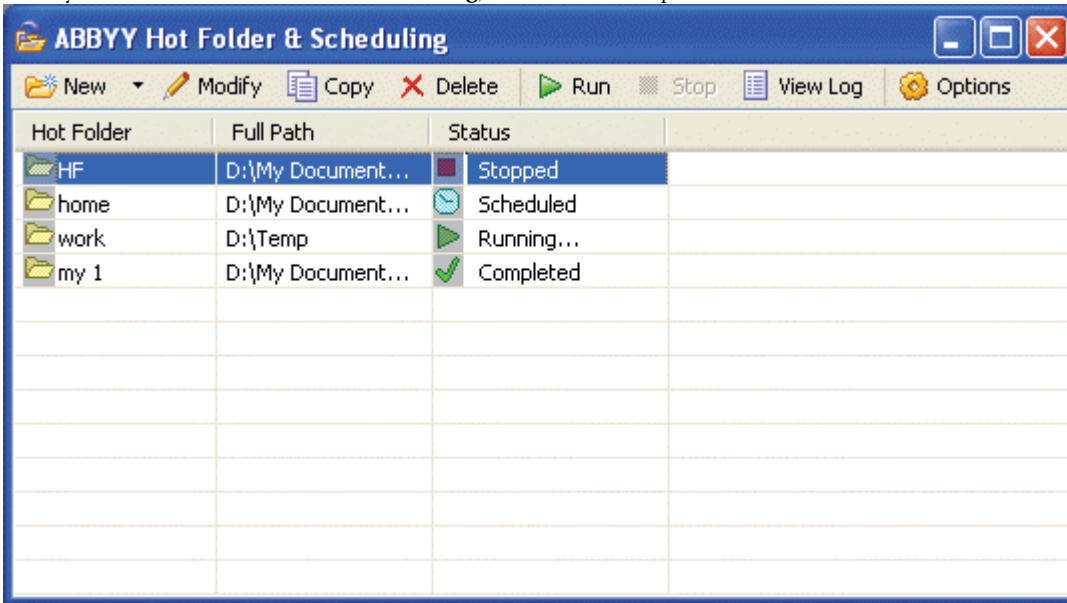
or

- Double-click the **ABBYY Hot Folder & Scheduling** icon on the Windows taskbar (this icon will appear after you set up at least one hot folder).

Note: If you cannot start a task for a selected hot folder, make sure that the Task Scheduler Service is turned on or contact your system administrator.






ABBYY Hot Folder & Scheduling Main Window

When you run **ABBYY Hot Folder & Scheduling**, its main window opens.










The **ABBYY Hot Folder & Scheduling** toolbar contains buttons for setting up hot folders tasks and viewing processing logs.

Button	Description
New	Runs the ABBYY Hot Folder & Scheduling wizard.
Export...	Exports a task file. Exported task files have the extension *.hft and can be passed on to other users. In the dialog that opens, provide a name for the task file. Note: By default, ABBYY FineReader saves task files in %Userprofile%\Local Settings\Application Data\ABBYY\HotFolder\8.00.
Import...	Imports a task file. In the dialog that opens, provide the path to the task file you wish to import. Note: For ABBYY FineReader to be able to run an imported task, there must be a folder on your computer (or a network folder) which is specified as a hot folder in the task, and all the required recognition languages must be installed.
Modify	Modifies a task.
Copy	Copies a task. The copy will be added to the list of tasks immediately below the original task and will have Paused status.

	Delete	Deletes a task.
	Run	Starts processing the documents.
	Stop	Stops a task.
	View Log	Opens a log file for the selected folder which contains information about all the processing events.
	Options	Additional ABBYY Hot Folder & Scheduling options.

The **ABBYY Hot Folder & Scheduling** window displays a list of tasks. For each task, the full path to the hot folder is displayed, its current status and the scheduled processing time.

Task Statuses

Status	Description
 Running	The images in the folder are being processed.
 Waiting	Waiting for the previous task to finish.
 Scheduled	You selected to check the hot folder for images only once at start time. The start time is indicated in the Next Run Time column.
 Watching	ABBYY FineReader will process images in this folder as they come in.
 Stopped	Processing has been stopped by the user.
 Completed	Finished processing images in this folder.
 Error	An error occurred when processing images in this folder. ABBYY FineReader did not complete the processing tasks. To find out the cause of the error, select the failed folder and click the View Log button on the toolbar.

Setting Up a Hot Folder

Run **ABBYY Hot Folder & Scheduling** and click the **New** button on the toolbar. This will launch the ABBYY Hot Folder & Scheduling wizard which will guide you through the setup process.

Step 1. Open Images

Here you need to select the folder with images to be processed. Also at this step, you need to specify how often ABBYY FineReader must check this folder for new images.

Options available at step 1

1. In the drop-down list at the top of the dialog, select folder type:
 - Select **Open images from folder** if you wish ABBYY FineReader to process images from a local or a network folder.
Note: You may select an ABBYY FineReader batch in this field.
 - Select **Open images from FTP folder** if you wish ABBYY FineReader to process images from a folder on an FTP server. Enter the FTP login and password if the server requires authorization. Otherwise, select the **Anonymous login** option to use an anonymous connection.
Note: You may select an ABBYY FineReader batch in this field.
2. In the **Folder** field, provide the full path to the hot folder.
3. In the **Files of type** field, select the formats of image files to process.
4. Select **Split dual pages** to split the images of book pages into two separate batch pages. The recognition quality will increase, and you will be able to save each page to a separate file in this case.
5. Under **Time settings**, select one of the two available options:
 - Select **Check once at start time** to process only the images that will be available in this folder when the scheduled task begins. Any images that may be added to this folder after that will not be processed.
 - Select **Continually watch folder for incoming images** to monitor the folder and process new files as they arrive. ABBYY FineReader will check the folder for new images every 60 seconds.
Note. In this mode, output files with recognized text will be named as their source images.
6. In the **Start time** field, select the time when ABBYY FineReader must start processing images in the folder.
 - **Start now** starts the processing as soon as the hot folder has been set up.
 - **Start at specified time** allows you to select the start date and time.
 - **Paused** postpones the task.
7. Under **Processed image settings**, specify whether ABBYY FineReader should delete the images from the hot folder if the operation was successful.
 - Select **Do not remove from folder** to keep the images in the hot folder.
 - Select **Remove from folder** to delete the images after they have been processed.

Step 2. Read All Images

Here you need to select recognition options.

Options available at step 2

1. In the **Recognition language** drop-down list, select the language of the texts on the images.
Note: You can select more than one recognition language.
2. Under **Recognition mode**, select:
 - **Thorough** (in this mode, ABBYY FineReader will read even poor quality images), or
 - **Fast** (this mode is only recommended for images with good quality and simple layouts).
3. Under **PDF Recognition**, select:
 - **Extract text from PDF** (in this mode, ABBYY FineReader will extract text data from PDF files and use it for recognition), or
 - **Recognize PDF files as images** (this option is useful for PDF files with non-standard encoding. In this case text content can only be recovered with OCR).
4. In the **Hyperlinks** group:
 - Select the **Highlight hyperlinks** option if you wish ABBYY FineReader to highlight hyperlinks found in the recognized text with the color selected in the **Color** field. To select or change the color, click the **Color** button to the right of the option and select the color in the dialog that appears.
5. Under **Training**, select if you wish to use patterns to be used.
 - **Do not use user patterns.** This option is selected by default.
 - **Select Use user patterns** if you wish ABBYY FineReader to use user patterns for OCR. For details, see .
6. Under **Use block template**, provide the path to a block template if you wish ABBYY FineReader to use a block template when analyzing the layout of the images.

Step 3. Save Document

Here you need to select a folder where the resulting documents will be stored and select a file format for the documents.

Options available at step 3

1. Under **Specify output folder**, provide the path to the folder where the recognition results are to be stored.
2. Under Saving options:
 - In the **Save as type** drop-down list, select the file format in which you wish to save the results. Click the **Format Settings** button to the right to select additional options for the format.
3. If you have selected the **Check once at start time** mode at **Step 1**, select the following options in the **Specify output file(s)** group:
 - **File options**
 - Create a separate file for each page** saves each page into a separate file;
 - Name files as source images** saves each page into a separate file with the same name as that of the original image;
 - Create a new file at each blank page** treats the whole batch as a set of page groups, each group ending with a blank page. The pages from different groups are saved into different files with names that are formed from a user-specified name with an index appended: -1, -2, -3 and so on;
 - Create a single file for all pages** saves all pages to a single file.
 - In the **Name** field, provide a name for the file to which recognized text will be saved.
Note: If you have selected the **Continually watch folder for incoming images** mode, the **Specify output file(s)** group indicates that output files with recognized text will be named as source images as the file and naming options are not available in this mode.
4. Check the **Save batch** option if you wish to save the ABBYY FineReader batch. Click **Browse...** to select a folder where the batch is to be stored.

Note: The Save batch option is not available for the Continually watch folder for incoming images mode (see Step 1).

Tip. You can add several saving steps to save recognition results in multiple formats. In the **Save as type** drop-down list, click (**Select multiple formats**) and in the dialog that opens select the formats in which you wish to save recognition results.

Hot Folder Log File

Images in a hot folder are processed automatically and the entire process is logged.

The log file contains the following information:

- The name of the task and its settings
- Errors and warnings (if any)
- Statistics (number of processed pages, number of errors and warnings, number of uncertainly recognized characters)

Viewing Log Files

To view a log file:

1. In the **ABBYY Hot Folder & Scheduling main window**, select a task for which you wish to see the log.
2. Click the **View Log** button on the toolbar.

Additional Options for ABBYY Hot Folder & Scheduling

Click the **Options** button on the toolbar to select additional ABBYY Hot Folder & Scheduling options.

- Select **Store log file within output folder** if you wish the log file to be stored in the same folder where the recognition results are to be saved.
- Select **Notify when finished** if you wish to see a notification message when you rest the mouse cursor on the **ABBYY Hot Folder & Scheduling** icon on the Windows taskbar.

Appendix

Chapter Contents

- Supported Document Saving Formats
- Supported Image Formats
- Hot Keys
- Glossary

Supported Document Saving Formats

ABBYY FineReader saves recognition results in the following formats:

- Microsoft Word Document (*.DOC)
- Rich Text Format (*.RTF)
- Microsoft Word XML Document (*.XML) (Microsoft Office Professional Edition 2003 only)
- Adobe Acrobat Format (*.PDF)
- Hypertext Markup Language (*.HTML)
- Microsoft PowerPoint Format (*.PPT)
- Comma Separated Values (*.CSV)
- Plain Text (*.TXT). ABBYY FineReader supports various code pages (Windows, DOS, Mac, ISO) and Unicode encoding
- Microsoft Excel Spreadsheet (*.XLS)
- Database Format (*.DBF)
- Microsoft Reader eBook (*.LIT)

Supported Image Formats

ABBYY FineReader opens image files in the following formats:

PDF:

Files in PDF format (Version 1.6 or earlier)

BMP:

2-bit – black and white

4- and 8-bit – Palette

16-bit – Mask

24-bit – Palette and TrueColor

32-bit – Mask

PCX, DCX:

2-bit – black and white

4- and 8-bit – Palette

24-bit – TrueColor

JPEG/JPEG 2000:

gray, color

TIFF:

black and white – uncompressed, CCITT3, CCITT3FAX, CCITT4, Packbits, ZIP, LZW

gray – uncompressed, Packbits, JPEG, ZIP, LZW

RGB – uncompressed, JPEG, ZIP, LZW

CMYK – uncompressed, JPEG, ZIP, LZW

Palette – uncompressed, Packbits, ZIP, LZW

multi-page TIFF

PNG:

black and white, gray, color

ABBYY FineReader saves image files in the following formats:

BMP:

black and white, gray, color

PCX:

black and white, gray

JPEG/JPEG 2000:

gray, color

TIFF:

black and white – uncompressed, CCITT3, CCITT4, Packbits, ZIP

gray – uncompressed, Packbits, JPEG, ZIP

RGB – uncompressed, JPEG, ZIP

multi-page TIFF

PNG:

black and white, gray, color

Hot Keys

Menu	Command	Press:
File	Open an image from a file	CTRL+O
	Scan an image	CTRL+K
	Stop scanning	Esc
	Create a new batch	CTRL+N
	Open a batch	CTRL+SHIFT+N
	Save pages	CTRL+S
	E-mail pages	CTRL+M
	Save an image to a file	CTRL+ALT+S
	E-mail images	CTRL+ALT+M
Edit	Undo the last action	CTRL+Z
	Redo the last action	CTRL+Enter
	Cut the selection and put it on the Clipboard	CTRL+X
	Copy the selection to the Clipboard	CTRL+INS or CTRL+C
	Paste the Clipboard contents	CTRL+V or SHIFT+INS
	Select all text in the Text window, select all batch pages, or select all blocks on the open image	CTRL+A
	Find the specified text	CTRL+F
	Find the next occurrence of the search text	F3
	Search for and replace the specified text	CTRL+H
Advanced Search	ALT+F3	
View	Maximize the Batch window	CTRL+0
	Show the Image window	CTRL+F2
	Magnify the image in the Image window	CTRL+SHIFT+NUM+
	Zoom out the image in the Image window	CTRL+SHIFT+NUM-
	Zoom in on selected blocks	CTRL+SHIFT+NUM*
	Show the Text window	CTRL+F3
	Show the Zoom window	CTRL+F5
	Properties	ALT+ENTER
Batch	Open next batch page	CTRL+NUM+
	Open previous batch page	CTRL+NUM-
	Open page with specified number	CTRL+G
	Close the current page	CTRL+F4
Image	Change the block type to Recognition area	CTRL+1
	Change the block type to Text	CTRL+2
	Change the block type to Table	CTRL+3
	Change the block type to Picture	CTRL+4
	Change the block type to Barcode	CTRL+5
	Delete all blocks in the Image window and all recognized text in the Text window	CTRL+Del
	Delete all blocks and the recognized text in the Text window	CTRL+SHIFT+Del
	Split image in several parts	CTRL+SHIFT+I
Crop the unwanted edge areas from the image	CTRL+SHIFT+C	
Correct image resolution	CTRL+SHIFT+T	
Process	Scan and read an image	CTRL+D

	Open and read an image	CTRL+SHIFT+D
	Open an image from a file	CTRL+O
	Scan an image	CTRL+K
	Stop scanning	Esc
	Read active or selected pages	CTRL+R
	Read all batch pages	CTRL+SHIFT+R
	Read active or selected blocks	CTRL+SHIFT+B
	Analyze layout	Ctrl+E
	Analyze layout on all batch pages	CTRL+SHIFT+E
	Spell the recognized text	F7
	Move to the next error or uncertain word	F4
	Move to the previous error or uncertain word	SHIFT+F4
	Save pages	CTRL+S
	E-mail pages	CTRL+M
	Save an image to a file	CTRL+ALT+S
	E-mail images	CTRL+ALT+M
	Open the Automation Manager dialog	CTRL+T
Tools	Opens the Automation Manager	CTRL+T
	Start Hot Folder & Scheduling	CTRL+SHIFT+H
	Spell check the recognized text	F7
	View Dictionaries	CTRL+ALT+D
	Open the Language Editor dialog to create and edit languages and language groups	CTRL+SHIFT+L
	Open the Pattern Editor dialog to create and edit the user's patterns	CTRL+SHIFT+A
	Open the Options dialog	CTRL+SHIFT+O
Window	Open the next window	CTRL+F6
	Open the previous window	CTRL+SHIFT+F6
	Open the Batch window	ALT+0
	Open the Image window	ALT+2
	Open the Text window	ALT+3
	Open the Zoom window	ALT+4
	Switch to the Advanced search window	ALT+5
	Open the Advanced search window	ALT+F3
Help	Open Help	F1
General	Make the selection bold	CTRL+B
	Make the selection <i>italic</i>	CTRL+I
	<u>Underline</u> the selection	CTRL+U
	Go to the next table cell	left arrow, right arrow, up arrow, down arrow

Glossary

A

Abbreviation – A shortened form of a word or phrase used to represent the whole. For example, MS-DOS (for Microsoft Disk Operating System), UN (for United Nations), etc.

ABBYY Hot Folder & Scheduling – A scheduling agent which allows you to select a folder with images and set the time for processing images in this folder. The images from the selected folder will be processed automatically at the specified time.

ABBYY Screenshot Reader – A application which allows you to create screenshots and recognize texts on them.

Activation – The process of obtaining a special code from ABBYY which allows the user to use his copy of the software in full-function mode on a given computer.

Activation Code – A code that is issued by ABBYY to each user of ABBYY FineReader Professional Edition during the activation procedure. The Activation Code is required to activate ABBYY FineReader on the computer that generated the Installation ID.

Activation File – A file issued by ABBYY to each user of ABBYY FineReader Corporate Edition during the activation procedure. The Activation File contains information required to activate the software on the server or on a standalone computer as the case may be. From the server, the product will be activated on workstations.

Active block – Indicates the block that is currently ready to have actions (e.g. deleting, changing type, etc.) applied to it. The active block frame is bold and there are "squares" in its corners.

ADF (Automatic Document Feeder) – A device that automatically feeds documents through a scanner. A scanner with an ADF can scan any number of pages without manual intervention. ABBYY FineReader also supports scanning multiple images.

Automated task – A sequence of steps, each corresponding to a particular processing routine. ABBYY FineReader 8.0 includes three automated tasks which are ready for use and do not require additional tweaking. You can also create your own automated tasks to suit your needs. Automated tasks are launched from the menu of the **Scan&Read** button.

Automation Manager – A built-in manager which allows you to run an automated task, to create and modify automated tasks, and delete custom automated tasks which you no longer use.

B

Background recognition – A special recognition mode that allows the user to edit and save already recognized pages while ABBYY FineReader recognizes other pages.

Barcode – A block that is used for barcode image areas.

Batch – A folder that contains image files, recognized text files and other ABBYY FineReader information files. There may be up to 9,999 pages in a batch. It is useful to save similar pages (such as all pages from the same book, those in the same language, or images with the same layout) in the same batch to streamline the work process.

Block – a framed image area.

Block type – Each block has a type. The following block types are available in ABBYY FineReader: Recognition Area, Text, Picture, Table and Barcode.

Blocks template – a description of blocks sizes and location on the page. A particular blocks template can be used to recognize pages of similar layout.

Brightness – A scanning parameter that indicates the contrast between black and white image areas. Setting correct brightness increases recognition quality.

Brightness autotuning – Automatic brightness tuning performed either by the scanner or by ABBYY FineReader. The autotuning process sets the brightness for every image area separately.

C

Code page – A table that sets the interrelation between the character codes and the characters themselves. Users can select the characters they need from the set found in the code page.

Compound word – A word made up of two or more stems (general meaning); a word not found in the dictionary, but potentially made up of two or more terms found in the dictionary (ABBYY FineReader meaning).

D

Despeckle image – Delete excess small black dots from an image.

Document Open password – A password which prevents the users from opening a PDF document unless they type the password the author specified.

Document properties – Properties which are assigned to a document and allow the user to sort or find files. These properties contain the title of the document, the name of its author, its subject and keywords.

dpi (Dots per Inch) – How resolution is measured.

Driver – A program controlling a computer peripheral (e.g., a scanner, a monitor, etc).

F

Font effects – Certain variations of a font outlook (i.e. bold, italic, underlined, strikethrough, subscript, superscript, small caps).

I

Ignored characters – Any non-letter characters found in words (e.g. syllable characters or stress marks). These characters are ignored during the spell check.

Image type – A scanning parameter that determines whether an image must be scanned in black and white, gray or color mode.

Installation ID – A computer code that is generated on the basis of the PC hardware parameters.

Inverted image – An image with white characters against a dark background.

L

License Manager – A utility used for managing ABBYY FineReader licenses and activating ABBYY FineReader 8.0 Corporate Edition.

Ligature – A combination of two or more "glued" characters, for example, fi, fl, ffi, etc. These characters are difficult to separate because they are usually "glued" in print. Treating them as a single compound character improves scanning accuracy.

M

Monospaced font – A font (such as Courier New) in which all characters are equally spaced. Select the **Typewriter** item on the **Print Type** group (**Recognition** tab) to increase the recognition quality of documents set in monospaced fonts.

O

Omnifont system – A recognition system that recognizes characters set in any font and font size without prior training.

Option set – the total of option values specified on the **Scan/Open**, **Read**, **Check Spelling** and **Save** tabs of the **Options**, **Formats Settings** and Legacy Options dialog boxes. Option sets also include user languages and patterns. Option sets can be saved and then used (loaded) in other ABBYY FineReader batches.

Optional hyphen – A hyphen (–) that indicates exactly where a word or word combination should be split if it occurs at the end of a line (e.g. "autoformat" should be split to "auto–format"). ABBYY FineReader replaces all hyphens found in dictionary words with optional hyphens.

P

Page layout – A combination of the way text, tables and pictures are arranged on a page, the way text is arranged into paragraphs, the font and font size of the text, the number of text columns, the character and background color, and the text orientation.

Page layout analysis (drawing blocks) – A process of analyzing the page layout and enclosing different image areas in blocks according to the layout. Blocks may be of different types. Page layout analysis may be performed automatically in a coupled recognition/page layout analysis procedure (run by clicking the **2–Read** button) or manually.

Paradigm – The set of all grammatical forms of a word.

Pattern – A set of pairs (the character image and the character itself) that is created during pattern training. A pattern provides additional information during recognition.

Permissions password – A password which prevents the users from printing and editing a PDF document unless they type the password the author specified. If some security settings are selected for the document the users will not be able to change these settings until they type the password you specify.

PDF security settings – Restrictions that can prevent a PDF document from being opened, edited, copied or printed. These settings include Document Open passwords, Permissions passwords and encryption levels.

Picture – A block that is used for image areas that contain pictures. This type of block may enclose an actual picture or any other object that should be displayed as a picture (e.g. a section of text).

Primary form – The form of words in dictionary.

Prohibited characters – If certain characters will never be found in recognized text, they may be specified in a set of prohibited characters in the language group properties. Specifying these characters increases the speed and quality of recognition. To specify a set of prohibited characters, click on **Advanced** in the **Language Group Properties** dialog. The **Advanced language group properties** dialog will open. Specify the set of prohibited characters in the **Prohibited characters** line.

R

Recognition Area – A block that is used for automatic recognition and analysis. After the **2–Read** button is clicked, all blocks of this type will be automatically analyzed and recognized.

Resolution – A scanning parameter determining how many dpi to use during scanning. Resolution of 300 dpi should be used for texts set in 10pt font size and larger, 400 to 600 dpi is preferable for texts of smaller font size (9pt and less).

S

Scanner – A device for inputting images into computer.

Scan&Read – The main ABBYY FineReader button. Click it to have ABBYY FineReader scan and recognize your image(s).

Scan&Read Wizard – Runs a special Scan&Read mode. ABBYY FineReader guides you through document processing and provides advice on getting best results.

Source Text Print Type – A parameter reflecting how the source text was printed (on a laser printer or equivalent, on a matrix printer in the draft mode, on a typewriter). For laser–printed texts, the **Auto** mode should be set; for typewritten texts, the **Typewriter** mode should be set; for texts printed on a dot matrix printer in draft mode, the **Dot Matrix Printer** mode should be set.

T

Table – A block that is used for table image areas or for areas of text that are structured in a table. When the application reads this type of block, it draws vertical and horizontal separators inside the block to form a table. This block is represented as a table in the output text.

Tagged PDF – A PDF document which contains information about the document structure such as logical parts, pictures, tables, etc. This structure is expressed via "PDF tags". A PDF file equipped with the tags may be reflowed to fit different screen sizes and will be displayed well on handheld devices.

Text – A block that contains text areas. Note that Text blocks should only contain single–column text.

Training – Creating pairs of a character image and the character itself. See the "Recognition with Training" section for details.

TWAIN, TWAIN dialog – A scanner dialog.

U

Uncertain characters – Characters recognized with a certain degree of uncertainty. ABBYY FineReader marks characters that may be incorrectly recognized.

Uncertain words – Words containing one or several uncertain characters.

Unicode – A standard developed by The Unicode Consortium (Unicode, Inc.). The standard is a 16–bit international encoding system for processing texts written in the main world languages. The standard is easily extended. The Unicode Standard determines the character encoding, as well as properties and procedures used in processing texts written in a certain language.