

SAMS
Teach
Yourself

Regular Expressions

in **10**

Minutes

Ben Forta

SAMS

800 East 96th Street, Indianapolis, Indiana, 46240 USA

Sams Teach Yourself Regular Expressions in 10 Minutes

Copyright © 2004 by Sams Publishing

All rights reserved. No part of this book shall be reproduced, stored in a retrieval system, or transmitted by any means, electronic, mechanical, photocopying, recording, or otherwise, without written permission from the publisher. No patent liability is assumed with respect to the use of the information contained herein. Although every precaution has been taken in the preparation of this book, the publisher and author assume no responsibility for errors or omissions. Nor is any liability assumed for damages resulting from the use of the information contained herein.

International Standard Book Number: 0-672-32566-7

Library of Congress Catalog Card Number: 2003093136

Printed in the United States of America

First Printing: February 2004

06 05 04 03 4 3 2 1

Trademarks

All terms mentioned in this book that are known to be trademarks or service marks have been appropriately capitalized. Sams Publishing cannot attest to the accuracy of this information. Use of a term in this book should not be regarded as affecting the validity of any trademark or service mark.

Warning and Disclaimer

Every effort has been made to make this book as complete and as accurate as possible, but no warranty or fitness is implied. The information provided is on an “as is” basis. The author and the publisher shall have neither liability nor responsibility to any person or entity with respect to any loss or damages arising from the information contained in this book or from the use of the CD or programs accompanying it.

Bulk Sales

Sams Publishing offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact

U.S. Corporate and Government Sales

1-800-382-3419

corpsales@pearsontechgroup.com

For sales outside of the U.S., please contact

International Sales

1-317-428-3341

international@pearsontechgroup.com

ASSOCIATE PUBLISHER

Michael Stephens

DEVELOPMENT EDITOR

Mark Renfrow

MANAGING EDITOR

Charlotte Clapp

PRODUCTION EDITOR

Sheila Schroeder

COPY EDITOR

Barbara Hacha

INDEXER

Ken Johnson

PROOFREADER

Juli Cook

TECHNICAL EDITOR

Michael Dinowitz

TEAM COORDINATOR

Cindy Teeters

INTERIOR AND COVER DESIGNER

Gary Adair

Sams Teach Yourself Regular Expressions in 10 Minutes

Table of Contents

Sams Teach Yourself Regular Expressions in 10 Minutes

Copyright © 2004 by Sams Publishing

Table of Contents

About the Author

We Want to Hear from You!

Introduction

Who Is Sams Teach Yourself Regular Expressions For?

LESSON 1 Introducing Regular Expressions

Understanding the Need

How Regular Expressions Are Used

So What Exactly Is a Regular Expression?

Using Regular Expressions

Before You Get Started

Summary

LESSON 2 Matching Single Characters

Table of Contents

Matching Literal Text
Matching Any Characters
Matching Special Characters
Summary

LESSON 3 Matching Sets of Characters

Matching One of Several Characters
Using Character Set Ranges
Anything But Matching
Summary

LESSON 4 Using Metacharacters

Escaping Revisited
Matching Whitespace Characters
Matching Specific Character Types
Using POSIX Character Classes
Summary

LESSON 5 Repeating Matches

How Many Matches?
Using Intervals
Preventing Over Matching
Summary

LESSON 6 Position Matching

Table of Contents

Using Boundaries

Using Word Boundaries

Defining String Boundaries

Summary

LESSON 7 Using Subexpressions

Understanding Subexpressions

Grouping with Subexpressions

Nesting Subexpressions

Summary

LESSON 8 Using Backreferences

Understanding Backreferences

Matching with Backreferences

Performing Replace Operations

Summary

LESSON 9 Looking Ahead and Behind

Introducing Lookaround

Looking Ahead

Looking Behind

Combining Lookahead and Lookbehind

Negating Lookaround

Summary

Table of Contents

LESSON 10 Embedding Conditions

Why Embed Conditions?

Using Conditions

Summary

APPENDIX A Regular Expressions in Popular Applications and Languages

grep

Macromedia Dreamweaver

Macromedia HomeSite (and ColdFusion Studio)

Microsoft ASP

Microsoft ASP.NET

Microsoft C#

Microsoft .NET

Microsoft Visual Studio .NET

MySQL

Perl

PHP

Sun Java

APPENDIX B Regular Expression Solutions to Common Problems

North American Phone Numbers

Table of Contents

U.S. ZIP Codes

Canadian Postal Codes

United Kingdom Postcodes

U.S. Social Security Numbers

IP Addresses

URLs

Complete URLs

Email Addresses

HTML Comments

JavaScript Comments

Credit Card Numbers

Summary

APPENDIX C The Regular Expression Tester

Understanding the Regular Expression Tester
Application

Obtaining a Copy of the Application

INDEX