

Cambridge University Press
978-1-107-00068-1 — UNIX and Perl to the Rescue!
Keith Bradnam, Ian Korf
Frontmatter
[More Information](#)

Unix and Perl to the Rescue!

Your research has generated gigabytes of data and now you need to analyze it. You hate using spreadsheets, but it's all you know, so what else can you do? This book will transform how you work with large and complex data sets, teaching you powerful programming tools for slicing and dicing data to suit your needs.

Written in a fun and accessible style, this step-by-step guide will inspire and inform non-programmers about the essential aspects of Unix and Perl. It shows how, with just a little programming knowledge, you can write programs that could save you hours, or even days. No prior experience is required, and new concepts are introduced using numerous code examples that you can try out for yourself. Going beyond the basics, the authors touch upon many broader topics that will help those new to programming, including debugging and how to write in a good programming style.

KEITH BRADNAM is a project scientist in the Genome Center at the University of California, Davis. He has extensive experience working with model organism databases and spent four years as a project leader at WormBase, helping to develop this important bioinformatics resource.

IAN KORF is an Associate Professor in Molecular and Cellular Biology at the University of California, Davis. His research seeks to understand structure and function in genomic DNA. He has developed new tools for gene prediction, co-authored the only book devoted to BLAST and helped in the development of BioPerl.

Cambridge University Press
978-1-107-00068-1 — UNIX and Perl to the Rescue!
Keith Bradnam, Ian Korf
Frontmatter
[More Information](#)

Unix and Perl to the Rescue!

A field guide for the life sciences (and other data-rich pursuits)

KEITH BRADNAM

University of California, Davis

IAN KORF

University of California, Davis



Cambridge University Press
978-1-107-00068-1 — UNIX and Perl to the Rescue!
Keith Bradnam, Ian Korf
Frontmatter
[More Information](#)

CAMBRIDGE UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom
One Liberty Plaza, 20th Floor, New York, NY 10006, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
314-321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi - 110025, India
103 Penang Road, #05-06/07, Visioncrest Commercial, Singapore 238467

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning and research at the highest international levels of excellence.

www.cambridge.org

Information on this title: www.cambridge.org/9781107000681

© Keith Bradnam and Ian Korf 2012

This publication is in copyright. Subject to statutory exception and to the provisions of relevant collective licensing agreements, no reproduction of any part may take place without the written permission of Cambridge University Press.

First published 2012
Reprinted 2014

A catalogue record for this publication is available from the British Library

Library of Congress Cataloging in Publication data

Bradnam, Keith.

UNIX and Perl to the rescue! : a field guide for the life sciences (and other data-rich pursuits) / Keith Bradnam, Ian Korf.

p. cm.

Includes bibliographical references and index.

ISBN 978-1-107-00068-1 (hardback) – ISBN 978-0-521-16982-0 (paperback)

1. Science – Data processing. 2. UNIX (Computer file) 3. Perl (Computer program language)

I. Korf, Ian. II. Title.

Q183.9.B73 2012

005.4'32–dc23 2011047927

ISBN 978-1-107-00068-1 Hardback

ISBN 978-0-521-16982-0 Paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party internet websites referred to in this publication, and does not guarantee that any content on such websites is, or will remain, accurate or appropriate.

Contents

Part 1 Introduction and background	<i>page</i> 1
1.1 Introduction	3
1.2 How to use this book	6
Part 2 Installing Unix and Perl	9
2.1 What do I need in order to learn Unix and Perl?	11
2.2 Installing Linux on a PC	13
2.3 Installing a code editor	15
Part 3 Essential Unix	17
3.1 Introduction to Unix	19
3.2 The Unix terminal	20
3.3 The Unix command prompt	22
3.4 Your first Unix command	24
3.5 The hierarchy of a Unix filesystem	26
3.6 Finding out where you are in the filesystem	28
3.7 How to navigate a Unix filesystem	29
3.8 Absolute and relative paths	32
3.9 Working with your home directory	34
3.10 The Unix shell	36
3.11 Environment variables	38
3.12 Introduction to command-line options	40
3.13 Man pages	44
3.14 Working with directories	47
3.15 The importance of saving keystrokes	50
3.16 Moving and renaming files	55
3.17 Moving and renaming directories	59
3.18 How to remove files	61
3.19 How to copy files and directories	63
3.20 Working with text files	68
3.21 Introduction to aliases	71
3.22 Editing text files	76
3.23 Automating Unix commands	80
3.24 How to hide files and find hidden files	83
3.25 Creating a configuration file	86
3.26 Programming with Unix	91
3.27 Unix file permissions	93
3.28 How to specify which directories contain programs	96
3.29 Creating useful shell scripts	100
3.30 Unix summary	103

Part 4 Essential Perl	105
4.1 Hello world	107
4.2 Scalar variables	110
4.3 Use warnings	114
4.4 Maths and functions	117
4.5 Perl vs. perl	123
4.6 Conditional statements	125
4.7 Use strict	132
4.8 Stopping programs	138
4.9 Working with strings	141
4.10 Dealing with special characters	147
4.11 Matching operators	149
4.12 The transliteration operator	156
4.13 List context	160
4.14 Introduction to Arrays	163
4.15 Array manipulation	169
4.16 The @ARGV array	176
4.17 Defined and undefined variables	180
4.18 Sorting	183
4.19 Introduction to loops	187
4.20 More loops	192
4.21 Loop control	197
4.22 Data input and output	202
4.23 Reading and writing files	207
4.24 Introduction to hashes	214
4.25 Working with hashes	220
4.26 Introduction to regular expressions	225
4.27 Regular expression metacharacters	227
4.28 Working with regular expressions	235
4.29 Interacting with other programs	241
4.30 Using functions and subroutines	246
4.31 Returning data from a subroutine	251
4.32 Part 4 summary	255
Part 5 Advanced Unix	257
5.1 Introduction to advanced Unix	259
5.2 Introduction to process control	263
5.3 The grep command	269
5.4 Viewing and controlling program output	274
5.5 Redirecting input and output	276
5.6 Standard error	282
5.7 Connecting commands with pipelines	285
5.8 Advanced text manipulation	289

Part 6 Advanced Perl	293
6.1 Regular expressions revisited	295
6.2 Function libraries	300
6.3 References and two-dimensional arrays	306
6.4 Records and other hash references	314
6.5 Using references with subroutines	319
6.6 Complex data structures	321
6.7 Adding command-line options	328
6.8 OOP basics	333
6.9 CPAN	340
Part 7 Programming topics	343
7.1 Debugging strategies	345
7.2 Common error messages	351
7.3 Code beautification	355
7.4 Abstraction	361
7.5 Data management	368
7.6 Documentation	377
7.7 Revision control	382
7.8 Working with other people's data	385
7.9 Getting help	389
<i>Appendix</i>	393
<i>Index</i>	410