

# ГОЛОВОЛОМКИ

на

# PHR

ДЛЯ

# КАКЕРА

Максим Кузнецов  
Игорь Симдянов

2-е издание

+CD



УДК 681.3.06  
ББК 32.973.26-018.2  
К89

**Кузнецов, М. В.**

К89 Головоломки на PHP для хакера / М. В. Кузнецов,  
И. В. Симдянов. — 2-е изд., перераб. и доп. — СПб.: БХВ-Петербург,  
2008. — 544 с.: ил. + CD-ROM

ISBN 978-5-9775-0204-7

Книга представляет собой задачник по Web-технологиям с уклоном в защиту Web-приложений от злоумышленников. Цель книги — помочь Web-разработчику научиться самостоятельно обнаруживать и устранять уязвимости в своем коде.

Главы второго издания существенно обновлены, кроме этого написаны две новые главы, посвященные динамическому формированию изображений и объектно-ориентированному программированию.

На компакт-диске, поставляемом вместе с книгой, приведены скрипты, являющиеся ответами на предлагаемые задачи.

*Для программистов и Web-разработчиков*

УДК 681.3.06  
ББК 32.973.26-018.2

**Группа подготовки издания:**

Главный редактор	<i>Екатерина Кондукова</i>
Зам. главного редактора	<i>Евгений Рыбаков</i>
Зав. редакцией	<i>Григорий Добин</i>
Редактор	<i>Ирина Иноземцева</i>
Компьютерная верстка	<i>Ольги Сергиенко</i>
Корректор	<i>Зинаида Дмитриева</i>
Оформление обложки	<i>Елены Беляевой</i>
Зав. производством	<i>Николай Тверских</i>

Лицензия ИД № 02429 от 24.07.00. Подписано в печать 19.12.07.

Формат 70×100<sup>1/16</sup>. Печать офсетная. Усл. печ. л. 43,86.

Тираж 2500 экз. Заказ №

"БХВ-Петербург", 194354, Санкт-Петербург, ул. Есенина, 5Б.

Санитарно-эпидемиологическое заключение на продукцию  
№ 77.99.02.953.Д.006421.11.04 от 11.11.2004 г. выдано Федеральной службой  
по надзору в сфере защиты прав потребителей и благополучия человека.

Отпечатано с готовых диапозитивов  
в ГУП "Типография "Наука"  
199034, Санкт-Петербург, 9 линия, 12

ISBN 978-5-9775-0204-7

© Кузнецов М. В., Симдянов И. В., 2008  
© Оформление, издательство "БХВ-Петербург", 2008

# Оглавление

<b>Введение</b> .....	<b>1</b>
Благодарности .....	2
<b>ЧАСТЬ I. ЗАДАЧИ</b> .....	<b>3</b>
<b>Глава I.1. Строки и числа</b> .....	<b>5</b>
I.1.1. Количество и имена файлов в произвольном каталоге.....	5
I.1.2. Вывод случайного количества символов .....	5
I.1.3. Выравнивание по правому краю.....	6
I.1.4. Выравнивание по левому и правому краям .....	7
I.1.5. Вывод данных в три столбца .....	7
I.1.6. Передача массива между двумя страницами.....	8
I.1.7. Передача массива методом GET .....	8
I.1.8. Передача массива методом POST .....	8
I.1.9. Передача массива через сессии .....	8
I.1.10. Передача массива через cookie .....	9
I.1.11. Вертикальный вывод строки.....	9
I.1.12. Число в денежном формате.....	9
I.1.13. Упаковка IP-адреса .....	9
I.1.14. Календарь .....	10
I.1.15. Замена символов bbCode.....	11
I.1.16. Преобразование десятичного числа в двоичное и обратно .....	11
I.1.17. Возведение числа в степень .....	12
I.1.18. Изменение регистра строки .....	12
I.1.19. Преобразование арабского числа в римское .....	12
<b>Глава I.2. Регулярные выражения</b> .....	<b>13</b>
I.2.1. Удаление всех тегов из HTML-страницы .....	13
I.2.2. Удаление изображений из HTML-страницы .....	14
I.2.3. Преобразование нескольких пробельных символов в один .....	14
I.2.4. Извлечение названия HTML-страницы.....	14
I.2.5. Конвертация даты из MySQL-формата в календарный формат.....	14
I.2.6. Проверка корректности ввода адреса электронной почты.....	14

I.2.7. Проверка корректности ввода URL.....	15
I.2.8. Подсветка URL.....	15
I.2.9. Проверка корректности ввода чисел.....	15
I.2.10. Изменение регистра.....	15
I.2.11. Разбивка длинной строки.....	16
I.2.12. Разбивка HTML-страницы на предложения.....	16
I.2.13. Количество слов в тексте.....	16
I.2.14. Интерпретация тегов bbCode.....	16
I.2.15. Подсветка PHP-кода.....	18
I.2.16. Замена подстроки с условием.....	18

### **Глава I.3. Файлы..... 19**

I.3.1. Загрузка файлов на сервер.....	19
I.3.2. Загрузка файла на сервер по частям.....	20
I.3.3. Создание файла с уникальным именем.....	20
I.3.4. Редактирование файлов на удаленном сервере.....	20
I.3.5. Уязвимость скрипта загрузки.....	20
I.3.6. Счетчик загрузок.....	21
I.3.7. Сохранение текстовых и графических файлов.....	21
I.3.8. Определение размера файла.....	23
I.3.9. Определение количества строк в файле.....	23
I.3.10. Изменение порядка следования строк в файле.....	23
I.3.11. Случайный вывод из файла.....	23
I.3.12. Редактирование файла.....	24
I.3.13. Сортировка содержимого текстового файла.....	24
I.3.14. Добавление записи в файл.....	24
I.3.15. Постраничная навигация для файла.....	24
I.3.16. Самая длинная и самая короткая строка в файле.....	25
I.3.17. Вывод из файла списка строк заданной длины.....	25
I.3.18. Вывод слов из файла по первым символам.....	25
I.3.19. Список файлов и подкаталогов в каталоге.....	25
I.3.20. Количество файлов в каталогах.....	25
I.3.21. Количество строк в файлах проекта.....	26
I.3.22. Замена строки во всех файлах вложенных подкаталогов.....	27
I.3.23. Копирование содержимого одного каталога в другой.....	27
I.3.24. Удаление каталога.....	27
I.3.25. Подсчет объема памяти, занимаемой каталогом.....	27
I.3.26. Система регистрации.....	27
I.3.27. Случайное изображение из каталога.....	28
I.3.28. Определение даты создания изображения.....	28
I.3.29. Взлом гостевой книги.....	28

### **Глава I.4. MySQL..... 30**

I.4.1. Система регистрации.....	30
I.4.2. SQL-инъекция по числовому параметру.....	32
I.4.3. Определение версии сервера MySQL.....	33

I.4.4. Поиск пользователя — SQL-инъекция.....	33
I.4.5. Удаление пользователей при помощи SQL-инъекции.....	35
I.4.6. Постраничная навигация.....	37
I.4.7. Алфавитная навигация.....	39
I.4.8. Сортировка.....	40
I.4.9. Двойной выпадающий список.....	40
I.4.10. Удаление сразу нескольких позиций.....	41
I.4.11. Хранение MP3-файлов в базе данных.....	41
I.4.12. Хранение изображений в базе данных.....	42
I.4.13. Загрузка данных из дампа базы данных.....	43
I.4.14. Шифрование содержимого базы данных.....	44
<b>Глава I.5. Протокол HTTP.....</b>	<b>45</b>
I.5.1. Загрузка страницы.....	45
I.5.2. Получение HTTP-заголовков с сервера.....	45
I.5.3. Определение размера файла на удаленном хосте.....	46
I.5.4. Отправка данных методом POST.....	46
I.5.5. Получение точного времени.....	47
<b>Глава I.6. Сессии и cookie.....</b>	<b>48</b>
I.6.1. Пользователи Online.....	49
I.6.2. Собственный механизм сессии.....	49
I.6.3. Защита HTML-формы при помощи сессии.....	49
I.6.4. Определение, включены ли cookie у посетителя.....	50
I.6.5. Фальсификация cookie.....	50
I.6.6. Обход защищенной сессией HTML-формы.....	51
I.6.7. Межсайтовый скриптинг.....	52
I.6.8. Похищение cookie.....	54
<b>Глава I.7. Пользовательские агенты и рефереры.....</b>	<b>55</b>
I.7.1. Переходы с других сайтов.....	55
I.7.2. Защита HTML-формы при помощи реферера.....	56
I.7.3. Фальсификация реферера.....	57
I.7.4. Ключевые слова поисковых систем.....	57
I.7.5. Распознавание посещений сайта роботами поисковых систем.....	57
I.7.6. Защита от менеджеров загрузки.....	57
I.7.7. Фальсификация пользовательского агента.....	58
<b>Глава I.8. Авторизация и аутентификация.....</b>	<b>59</b>
I.8.1. Авторизация на файлах.....	60
I.8.2. Шифрование пароля.....	61
I.8.3. Подбор пароля перебором.....	62
I.8.4. Подбор пароля по словарю.....	62
I.8.5. Генератор паролей.....	63
I.8.6. Защита текстовых файлов от просмотра в браузере.....	63

I.8.7. Авторизация при помощи cookie .....	64
I.8.8. Защита имени пользователя от подделки .....	66
I.8.9. Авторизация при помощи сессий .....	67
I.8.10. Шифрование пароля в базе данных .....	69
I.8.11. Базовая NTTP-авторизация .....	69
<b>Глава I.9. Использование информации со сторонних сайтов .....</b>	<b>70</b>
I.9.1. Загрузка страницы с удаленного хоста .....	71
I.9.2. Извлечение ссылок с Yandex .....	71
I.9.3. Извлечение ссылок с Google .....	72
I.9.4. Извлечение ссылок с Rambler .....	73
I.9.5. Извлечение ссылок с Aport .....	74
I.9.6. Определение курса валют из XML-файла .....	76
I.9.7. Определение динамики курса валют .....	77
I.9.8. Загрузка новостей со стороннего сайта .....	78
I.9.9. Создание новостного RSS-канала .....	78
<b>Глава I.10. FTP-протокол .....</b>	<b>80</b>
I.10.1. Определение типа операционной системы .....	80
I.10.2. Список файлов на FTP-сервере .....	80
I.10.3. Загрузка файлов .....	80
I.10.4. Изменение прав доступа .....	81
I.10.5. Какой объем памяти занимает сайт? .....	81
I.10.6. Поиск файлов, чей размер превышает 100 Кбайт .....	81
I.10.7. Перенос сайта с одного хоста на другой .....	81
<b>Глава I.11. Электронная почта .....</b>	<b>82</b>
I.11.1. Отправка почтового сообщения с сайта .....	82
I.11.2. Отправка письма с вложением .....	82
I.11.3. Массовая рассылка писем .....	82
I.11.4. Предотвращение массовой рассылки .....	83
I.11.5. Отправка почтового сообщения через SMTP-ретранслятор .....	83
I.11.6. Выяснение адресов почтовых ретрансляторов .....	83
I.11.7. Подсчет количества писем в почтовом ящике .....	83
I.11.8. Чтение заголовков писем .....	83
I.11.9. Удаление писем из почтового ящика .....	83
<b>Глава I.12. Whois-сервис .....</b>	<b>84</b>
I.12.1. Определение принадлежности IP-адресов .....	84
I.12.2. Определение принадлежности европейских IP-адресов .....	84
I.12.3. Следование реферальному серверу .....	85
I.12.4. Определение IP-адреса по сетевому адресу .....	86
I.12.5. Определение сетевого адреса по IP-адресу .....	86
I.12.6. Выяснение, занят ли домен .....	86

<b>Глава I.13. Объектно-ориентированное программирование .....</b>	<b>87</b>
I.13.1. Определение класса объекта .....	88
I.13.2. Счетчик объектов .....	89
I.13.3. Транзакции .....	89
I.13.4. Получение копии объекта .....	91
I.13.5. Хранение объекта в СУБД MySQL .....	92
I.13.6. Постраничная навигация .....	93
I.13.7. Создание исключений .....	94
I.13.8. Определение версии PHP и расширений .....	95
I.13.9. Распознавание загруженных расширений и их версий .....	95
<b>Глава I.14. Шпионские скрипты .....</b>	<b>96</b>
I.14.1. Слежение за ссылкой на удаленной странице .....	96
I.14.2. Проверка ссылочной целостности .....	96
I.14.3. Новые файлы на виртуальном хосте .....	97
I.14.4. Слишком большие файлы на виртуальном хосте .....	97
<b>Глава I.15. Динамические изображения (GDLib) .....</b>	<b>98</b>
I.15.1. Счетчик посещений .....	98
I.15.2. Несколько изображений на странице .....	98
I.15.3. Определение размера изображения .....	98
I.15.4. Защитное изображение для HTML-формы .....	99
I.15.5. Создание уменьшенной копии .....	99
I.15.6. Водяные знаки .....	100
I.15.7. Кривая Безье .....	100
I.15.8. Построение гистограммы .....	101
I.15.9. Построение круговой диаграммы .....	102
<b>Глава I.16. Разное .....</b>	<b>103</b>
I.16.1. Обмен значений переменных .....	103
I.16.2. Скрипт предзагрузки страницы .....	103
I.16.3. Использование утилиты ping .....	104
I.16.4. Работа с номером узла .....	104
I.16.5. Права доступа .....	104
I.16.6. Эмуляция утилиты tar .....	104
I.16.7. Буферизация данных .....	105
I.16.8. Размер страницы .....	105
I.16.9. Разгрузка баржи .....	106
I.16.10. Продолжительность жизни ученого .....	107
I.16.11. Выгода предпринимателя .....	107
<b>ЧАСТЬ II. РЕШЕНИЯ .....</b>	<b>109</b>
<b>Глава II.1. Строки и числа .....</b>	<b>111</b>
II.1.1. Количество и имена файлов в произвольном каталоге .....	111

П.1.2. Вывод случайного количества символов.....	114
П.1.3. Выравнивание по правому краю .....	115
П.1.4. Выравнивание по левому и правому краям.....	119
П.1.5. Вывод данных в три столбца .....	121
П.1.6. Передача массива между двумя страницами.....	123
П.1.7. Передача массива методом GET .....	123
П.1.8. Передача массива методом POST .....	125
П.1.9. Передача массива через сессии .....	126
П.1.10. Передача массива через cookie.....	128
П.1.11. Вертикальный вывод строки .....	129
П.1.12. Число в денежном формате .....	130
П.1.13. Упаковка IP-адреса.....	131
П.1.14. Календарь .....	132
П.1.15. Замена символов bbCode.....	135
П.1.16. Преобразование десятичного числа в двоичное и обратно.....	137
П.1.17. Возведение числа в степень .....	144
П.1.18. Изменение регистра строки .....	146
П.1.19. Преобразование арабского числа в римское .....	147
<b>Глава П.2. Регулярные выражения .....</b>	<b>151</b>
П.2.1. Удаление всех тегов из HTML-страницы .....	151
П.2.2. Удаление изображений из HTML-страницы .....	152
П.2.3. Преобразование нескольких пробельных символов в один.....	153
П.2.4. Извлечение названия HTML-страницы .....	154
П.2.5. Конвертация даты из MySQL-формата в календарный.....	155
П.2.6. Проверка корректности ввода адреса электронной почты .....	156
П.2.7. Проверка корректности ввода URL .....	157
П.2.8. Подсветка URL .....	158
П.2.9. Проверка корректности ввода чисел.....	159
П.2.10. Изменение регистра.....	160
П.2.11. Разбивка длинной строки.....	161
П.2.12. Разбивка текста на предложения .....	162
П.2.13. Количество слов в тексте .....	164
П.2.14. Интерпретация тегов bbCode.....	166
П.2.15. Подсветка РНР-кода .....	167
П.2.16. Замена подстроки с условием.....	170
<b>Глава П.3. Файлы .....</b>	<b>172</b>
П.3.1. Загрузка файлов на сервер .....	172
П.3.2. Загрузка файла на сервер по частям.....	174
П.3.3. Создание файла с уникальным именем .....	176
П.3.4. Редактирование файлов на удаленном сервере.....	177
П.3.5. Уязвимость скрипта загрузки .....	179
П.3.6. Счетчик загрузок .....	183
П.3.7. Сохранение текстовых и графических файлов.....	185



П.3.8. Определение размера файла .....	186
П.3.9. Определение количества строк в файле.....	188
П.3.10. Изменение порядка следования строк в файле .....	188
П.3.11. Случайный вывод из файла .....	189
П.3.12. Редактирование файла.....	190
П.3.13. Сортировка содержимого текстового файла .....	190
П.3.14. Добавление записи в файл .....	195
П.3.15. Постраничная навигация.....	196
П.3.16. Самая длинная и самая короткая строка в файле.....	198
П.3.17. Вывод из файла списка строк заданной длины.....	199
П.3.18. Вывод слов из файла по первым символам .....	200
П.3.19. Список файлов и подкаталогов в каталоге .....	203
П.3.20. Количество файлов в каталогах.....	204
П.3.21. Количество строк в файлах проекта.....	205
П.3.22. Замена строки во всех файлах вложенных подкаталогов .....	207
П.3.23. Копирование содержимого одного каталога в другой .....	208
П.3.24. Удаление каталога .....	209
П.3.25. Подсчет объема памяти, занимаемой каталогом .....	210
П.3.26. Система регистрации.....	211
П.3.27. Случайное изображение из каталога.....	216
П.3.28. Определение даты создания изображения.....	216
П.3.29. Взлом гостевой книги.....	217
<b>Глава П.4. MySQL и SQL-инъекции .....</b>	<b>220</b>
П.4.1. Система регистрации.....	220
П.4.2. SQL-инъекция по числовому параметру.....	223
П.4.3. Определение версии сервера MySQL .....	228
П.4.4. Поиск пользователя — SQL-инъекция .....	229
П.4.5. Удаление пользователей при помощи SQL-инъекции .....	234
П.4.6. Постраничная навигация.....	237
П.4.7. Алфавитная навигация .....	240
П.4.8. Сортировка .....	242
П.4.9. Двойной выпадающий список .....	244
П.4.10. Удаление сразу нескольких позиций .....	249
П.4.11. Хранение MP3-файлов в базе данных.....	251
П.4.12. Хранение изображений в базе данных.....	255
П.4.13. Загрузка данных из дампа базы данных .....	259
П.4.14. Шифрование содержимого базы данных.....	259
<b>Глава П.5. Протокол HTTP.....</b>	<b>264</b>
П.5.1. Загрузка страницы .....	264
П.5.2. Получение HTTP-заголовков с сервера.....	277
П.5.3. Определение размера файла на удаленном хосте .....	281
П.5.4. Отправка данных методом POST .....	282
П.5.5. Получение точного времени .....	284

<b>Глава II.6. Сессии и cookie .....</b>	<b>286</b>
П.6.1. Пользователи Online .....	286
П.6.2. Собственный механизм сессии.....	289
П.6.3. Защита HTML-формы при помощи сессии .....	294
П.6.4. Определение, включены ли cookie у посетителя .....	295
П.6.5. Фальсификация cookie .....	296
П.6.6. Обход защищенной сессией HTML-формы .....	298
П.6.7. Межсайтовый скриптинг.....	302
П.6.8. Похищение cookie.....	303
<b>Глава II.7. Пользовательские агенты и рефереры.....</b>	<b>305</b>
П.7.1. Переходы с других сайтов .....	305
П.7.2. Защита HTML-формы при помощи реферера.....	307
П.7.3. Фальсификация реферера .....	308
П.7.4. Ключевые слова поисковых систем .....	310
П.7.5. Распознавание посещений сайта роботами поисковых систем .....	311
П.7.6. Защита от менеджеров загрузки.....	313
П.7.7. Фальсификация пользовательского агента.....	313
<b>Глава II.8. Авторизация и аутентификация.....</b>	<b>315</b>
П.8.1. Авторизация на файлах .....	315
П.8.2. Шифрование пароля.....	320
П.8.3. Подбор пароля перебором .....	323
П.8.4. Подбор пароля по словарю .....	330
П.8.5. Генератор паролей .....	332
П.8.6. Защита текстовых файлов от просмотра в браузере.....	333
П.8.7. Авторизация при помощи cookie.....	333
П.8.8. Защита имени пользователя от подделки .....	340
П.8.9. Авторизация при помощи сессий.....	341
П.8.10. Шифрование пароля в базе данных .....	344
П.8.11. Базовая HTTP-авторизация.....	345
<b>Глава II.9. Использование информации со сторонних сайтов.....</b>	<b>348</b>
П.9.1. Загрузка страницы с удаленного хоста .....	348
П.9.2. Извлечение ссылок с Yandex .....	349
П.9.3. Извлечение ссылок с Google .....	351
П.9.4. Извлечение ссылок с Rambler.....	356
П.9.5. Извлечение ссылок с Aport .....	358
П.9.6. Определение курса валют из XML-файла .....	360
П.9.7. Определение динамики курса валют.....	362
П.9.8. Загрузка новостей со стороннего сайта .....	365
П.9.9. Создание новостного RSS-канала .....	367
<b>Глава II.10. FTP-протокол .....</b>	<b>372</b>
П.10.1. Определение типа операционной системы.....	372
П.10.2. Список файлов на FTP-сервере .....	374

П.10.3. Загрузка файлов .....	376
П.10.4. Изменение прав доступа .....	379
П.10.5. Какой объем памяти занимает сайт? .....	380
П.10.6. Поиск файлов, чей размер превышает 100 Кбайт.....	381
П.10.7. Перенос сайта с одного хоста на другой .....	383
<b>Глава П.11. Электронная почта.....</b>	<b>386</b>
П.11.1. Отправка почтового сообщения с сайта .....	386
П.11.2. Отправка письма с вложением .....	388
П.11.3. Массовая рассылка писем .....	391
П.11.4. Предотвращение массовой рассылки .....	393
П.11.5. Отправка почтового сообщения через SMTP-ретранслятор.....	395
П.11.6. Выяснение адресов почтовых ретрансляторов .....	396
П.11.7. Подсчет количества писем в почтовом ящике .....	397
П.11.8. Чтение заголовков писем .....	402
П.11.9. Удаление писем из почтового ящика.....	407
<b>Глава П.12. Whois-сервис .....</b>	<b>409</b>
П.12.1. Определение принадлежности IP-адресов.....	409
П.12.2. Определение принадлежности европейских IP-адресов .....	410
П.12.3. Следование реферальному серверу.....	411
П.12.4. Определение IP-адреса по сетевому адресу .....	414
П.12.5. Определение сетевого адреса по IP-адресу .....	415
П.12.6. Выяснение, занят ли домен.....	415
<b>Глава П.13. Объектно-ориентированное программирование .....</b>	<b>422</b>
П.13.1. Определение класса объекта.....	422
П.13.2. Счетчик объектов .....	424
П.13.3. Транзакции .....	427
П.13.4. Получение копии объекта .....	430
П.13.5. Хранение объекта в СУБД MySQL .....	431
П.13.6. Постраничная навигация.....	434
П.13.7. Создание исключений .....	455
П.13.8. Определение версии PHP и расширений .....	459
П.13.9. Распознавание загруженных расширений и их версии .....	459
<b>Глава П.14. Шпионские скрипты .....</b>	<b>461</b>
П.14.1. Слежение за ссылкой на удаленной странице .....	461
П.14.2. Проверка ссылочной целостности .....	468
П.14.3. Новые файлы на виртуальном хосте .....	471
П.14.4. Слишком большие файлы на виртуальном хосте .....	473
<b>Глава П.15. Динамические изображения (GDLib).....</b>	<b>475</b>
П.15.1. Счетчик посещений .....	475
П.15.2. Несколько изображений на странице.....	478

П.15.3. Определение размера изображения .....	480
П.15.4. Защитное изображение для HTML-формы .....	482
П.15.5. Создание уменьшенной копии.....	486
П.15.6. Водяные знаки .....	488
П.15.7. Кривая Безье.....	490
П.15.8. Построение гистограммы.....	493
П.15.9. Построение круговой диаграммы .....	495
<b>Глава П.16. Разное .....</b>	<b>497</b>
П.16.1. Обмен значений переменных .....	497
П.16.2. Скрипт предзагрузки страницы .....	497
П.16.3. Использование утилиты ping .....	499
П.16.4. Работа с номером узла.....	500
П.16.5. Права доступа .....	502
П.16.6. Эмуляция утилиты tar.....	506
П.16.7. Буферизация данных .....	509
П.16.8. Размер динамической страницы.....	510
П.16.9. Разгрузка баржи .....	510
П.16.10. Продолжительность жизни ученого.....	512
П.16.11. Выгода предпринимателя .....	513
<b>Заключение .....</b>	<b>515</b>
<b>ПРИЛОЖЕНИЯ.....</b>	<b>517</b>
<b>Приложение 1. Регулярные выражения.....</b>	<b>519</b>
П1.1. Синтаксис регулярных выражений .....	519
П1.2. Функции для работы с регулярными выражениями.....	523
П1.2.1. Функция <i>preg_grep()</i> .....	524
П1.2.2. Функция <i>preg_match()</i> .....	524
П1.2.3. Функция <i>preg_match_all()</i> .....	524
П1.2.4. Функция <i>preg_quote()</i> .....	526
П1.2.5. Функция <i>preg_replace()</i> .....	526
П1.2.6. Функция <i>preg_replace_callback()</i> .....	527
П1.2.7. Функция <i>preg_split()</i> .....	527
<b>Приложение 2. Описание компакт-диска .....</b>	<b>528</b>
<b>Предметный указатель .....</b>	<b>529</b>

# ГЛАВА I.1

```
## Sample if1.cfg fi
## Define preprocess
/DMY_PROJECT prepr
## Set extended leng
/4L132
## Set extended
## Set maximum float
/Op80
## Set maximum
## Additional direct
## files, before the
```

## Строки и числа

Работа со строками составляет основу любого программирования. Виртуозное манипулирование строками позволит программисту создавать более короткие и эффективные программы. Исследования показали, что плотность ошибок в программах не зависит от языка программирования, а зависит только от квалификации программиста. Чем короче будут программы, тем меньше ошибок и уязвимостей в них будет. Хорошее знание особенностей строк позволяет безошибочно определять возможные проблемные с точки зрения безопасности места в коде. Данная глава содержит задачи на знание строковых функций РНР и умение обращаться с ними.

### Замечание

Все примеры из данной главы можно найти в каталоге `scripts\1` компакт-диска, поставляемого вместе с книгой.

### I.1.1. Количество и имена файлов в произвольном каталоге

Определите количество и имена файлов в каталоге, не прибегая к функциям работы с каталогами (`opendir()`, `readdir()`, `closedir()` и т. п.). Решение задачи основано на том факте, что в РНР существует несколько видов кавычек, каждый из которых обладает своими свойствами.

### I.1.2. Вывод случайного количества символов

Создайте скрипт, который выводит случайное количество символов \* от 0 до 10.

### I.1.3. Выравнивание по правому краю

Пусть есть список файлов в массиве (листинг I.1.1). У имен файлов может быть различная длина, и необходимо выровнять их по правому краю так, как это изображено на рис. I.1.1. При решении задачи нельзя прибегать к атрибуту `align`, HTML-таблицам и каскадным таблицам стилей CSS, можно использовать только HTML-теги `<pre>` и `</pre>` и средства PHP.

#### Листинг I.1.1. Массив `$filename` с именами файлов

```
<?php
    $filename = array("all.php", "auth.php",
                     "auth.txt", "base.txt",
                     "chat.html", "config.php",
                     "count.txt", "count_new.txt",
                     "counter.dat", "counter.php",
                     "create.php", "dat.db");
?>
```

#### Замечание

Файл с массивом можно найти на прилагаемом к книге компакт-диске (`scripts\1\1.3\1.php`).

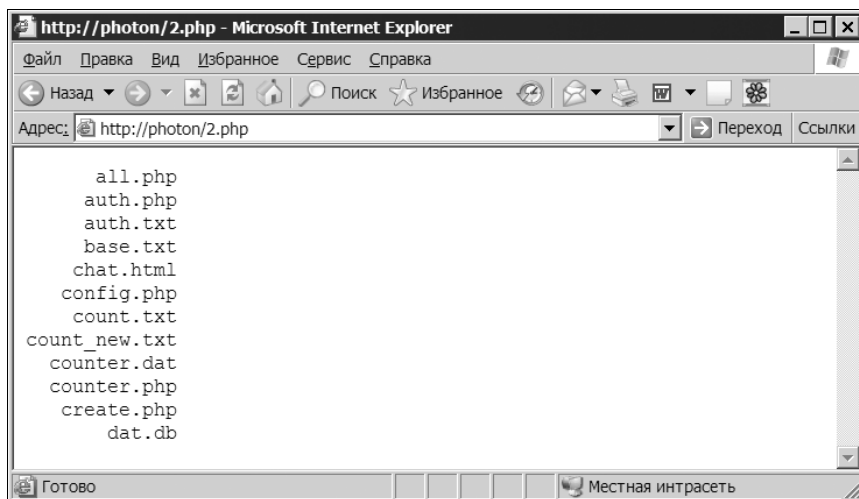


Рис. I.1.1. Выравнивание имен файлов по правому краю

## 1.1.4. Выравнивание по левому и правому краям

Необходимо разбить массив `$filename` (листинг 1.1.1) на две части и вывести в виде двух колонок так, как это представлено на рис. 1.1.2. При этом левая колонка должна быть выровнена по левому краю, а правая — по правому. При решении задачи нельзя прибегать к атрибуту `align`, HTML-таблицам и каскадным таблицам стилей CSS, можно использовать только HTML-теги `<pre>` и `</pre>` и средства PHP.

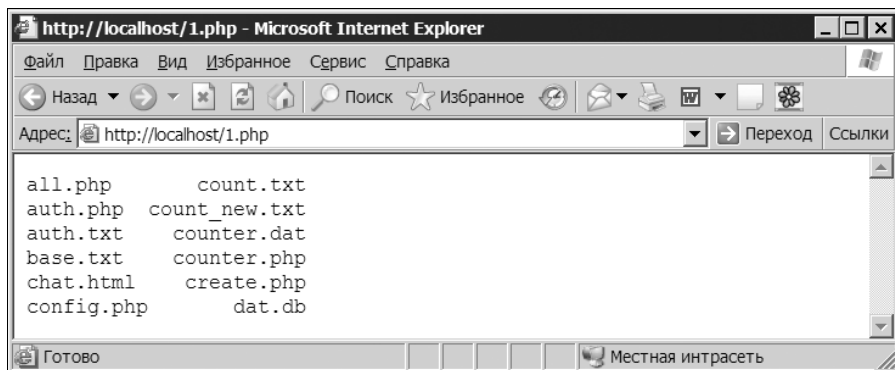


Рис. 1.1.2. Выравнивание имен файлов по левому и правому краям

## 1.1.5. Вывод данных в три столбца

Часто перед Web-разработчиками встает задача вывода таблицы, содержащей несколько столбцов. Выведите имена файлов из массива `$filename` (см. листинг 1.1.1) двумя способами:

- по строкам — сначала заполняется первая строка, затем вторая и т. д. (рис. 1.1.3);



Рис. 1.1.3. Первый вариант вывода массива в три столбца

- по столбцам — сначала заполняется первый столбец, затем второй и т. д. (рис. I.1.4).

При решении этой задачи необходимо динамически сформировать HTML-таблицу.



Рис. I.1.4. Второй вариант вывода массива в три столбца

## I.1.6. Передача массива между двумя страницами

Пусть массив `$filename`, представленный в листинге I.1.1, определен на странице `first.php`. Отобразите его на странице `second.php`, используя инструкцию `include`.

## I.1.7. Передача массива методом GET

Пусть массив `$filename`, представленный в листинге I.1.1, определен на странице `first.php`. Отобразите его на странице `second.php`, используя для передачи метод GET.

## I.1.8. Передача массива методом POST

Пусть массив `$filename`, представленный в листинге I.1.1, определен на странице `first.php`. Отобразите его на странице `second.php`, используя для передачи метод POST.

## I.1.9. Передача массива через сессии

Пусть массив `$filename`, представленный в листинге I.1.1, определен на странице `first.php`. Отобразите его на странице `second.php`, используя сессии.



## 1.1.10. Передача массива через cookie

Пусть массив `$filename`, представленный в листинге 1.1.1, определен на странице `first.php`. Отобразите его на странице `second.php`, используя `cookie`.

## 1.1.11. Вертикальный вывод строки

Выведите строку "Hello world!" вертикально, так, как это представлено на рис. 1.1.5.

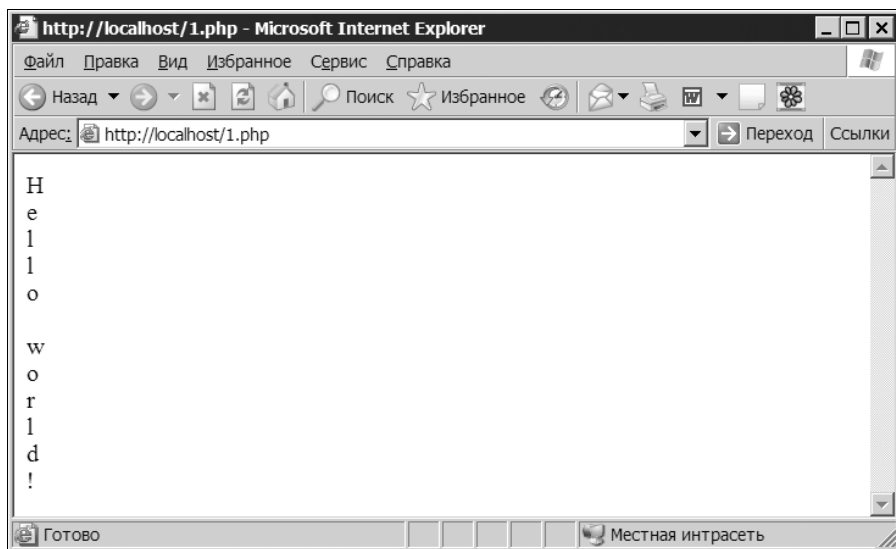


Рис. 1.1.5. Вертикальный вывод строки

## 1.1.12. Число в денежном формате

Пусть имеется число 18439529234.5678, его необходимо представить в денежном формате, т. е. чтобы после запятой осталось только два знака, а триады были бы разделены пробелом — 18 439 529 234.57.

## 1.1.13. Упаковка IP-адреса

Язык PHP обладает огромным количеством самых разнообразных функций. Среди них имеются функция для упаковки IP-адреса в целое число — `ip2long()` и функция обратного преобразования полученного целого числа в IP-адрес — `long2ip()`.

В листинге I.1.2 демонстрируется использование функций `ip2long()` и `long2ip()`.

### Листинг I.1.2. Использование функций `ip2long()` и `long2ip()`

```
<?php
// Получаем IP-адрес сайта http://www.softtime.ru/
$ip = gethostbyname('www.softtime.ru');
echo "$ip<br>";
// Преобразуем IP-адрес в число
$long = ip2long($ip);
echo "$long<br>";
// Преобразуем число в IP-адрес
$ip = long2ip($long);
echo "$ip<br>";
?>
```

Результатом работы скрипта из листинга I.1.2 будут следующие строки:

```
82.208.89.164
1389386148
82.208.89.164
```

Не прибегая к библиотечным функциям, создайте аналоги для функций `ip2long()` и `long2ip()`.

## I.1.14. Календарь

Создайте календарь на текущий месяц в двух форматах: американском (рис. I.1.6) и российском (рис. I.1.7).



Рис. I.1.6. Календарь на текущий месяц в американском формате

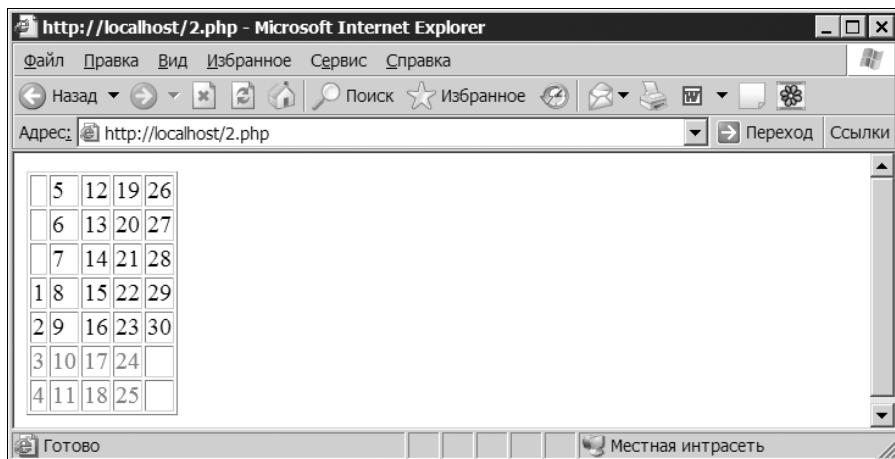


Рис. 1.1.7. Календарь на текущий месяц в российском формате

### Замечание

Данная задача переключается с задачей 1.1.5. Здесь также требуется вывод календаря по строкам и по столбцам.

Дни, соответствующие субботе и воскресенью, необходимо подсветить красным цветом.

## 1.1.15. Замена символов bbCode

Замените в тексте "Очень [b]жирный[/b], жирный [b]текст" символы bbCode [b] и [/b] на их HTML-эквиваленты `<b>` и `</b>`, не прибегая к регулярным выражениям. То есть для решения задачи должны быть использованы только строковые функции.

## 1.1.16. Преобразование десятичного числа в двоичное и обратно

Создайте два небольших Web-приложения, содержащих текстовую область и кнопку. Первое приложение при вводе десятичного числа должно выводить его двоичное представление, например, если пользователь ввел 482, то скрипт должен вывести 111100010. А второй скрипт должен решать обратную задачу — двоичное число переводить в десятичное.

Для решения задачи нельзя использовать никакие функции. Допустимы только арифметические операторы, циклы, операторы ветвления и конструкция `echo`.

**Замечание**

Использовать `print()` нельзя, т. к., в отличие от конструкции `echo`, это функция.

### I.1.17. Возведение числа в степень

Создайте программу, которая запрашивает у пользователя два целых числа и возводит первое число в степень второго, не используя функцию `pow()`.

**Замечание**

Следует воспользоваться битовыми операторами.

### I.1.18. Изменение регистра строки

Используя лишь поразрядные операторы, измените регистр каждого символа строки, введенной пользователем на строчный (прописной). Считается, что пользователь вводит только символы английского алфавита.

### I.1.19. Преобразование арабского числа в римское

Создайте скрипт, преобразующий число в арабской нотации (от 1 до 2000) в римское.

**Замечание**

Арабские числа соотносятся с римскими следующим образом: 1 — I, 5 — V, 10 — X, 50 — L, 100 — C, 500 — D, 1000 — M. Например, 116 == CXVI, 199 == CXCIX, 14 == XIV.

## ГЛАВА 1.2

```
## Sample if1.cfg file
## Define preprocess
/DMY_PROJECT prepr
## Set extended leng
/4L132
## Set extended
## Set maximum float
/Op80
## Set maximum
## Additional direct
## files, before the
```

# Регулярные выражения

Регулярные выражения являются специализированным мини-языком, используемым совместно со многими языками программирования. Сложную задачу можно решить двумя способами: либо создав сложное решение, используя простые технологии, либо создав простое решение, используя сложную технологию. Точно так же и с регулярными выражениями — изучить их достаточно сложно, но, поняв их один раз, далее в одну строку можно решать задачи, для решения которых при помощи строковых функций может понадобиться сотня строк. В *главе 1.1* было сказано, что плотность ошибок тем меньше, чем короче программа — регулярные выражения позволяют создавать не просто короткие программы, а очень короткие.

### Замечание

Синтаксис регулярных выражений, а также обзор функций для работы с ними представлены в *приложении 1*.

Сами по себе регулярные выражения являются достаточно специфическим языком программирования, аналога которому нет среди языков общего назначения. Поэтому их достаточно сложно изучить — знание других языков программирования фактически никак не облегчает задачу. Поэтому важно закреплять навыки при помощи практики.

### Замечание

Все примеры из данной главы можно найти в каталоге `scripts\2` компакт-диска, поставляемого вместе с книгой.

### 1.2.1. Удаление всех тегов из HTML-страницы

На компакт-диске найдите HTML-страницу `scripts\2\index.htm`. Прочитайте содержимое страницы и удалите все HTML-теги, оставив только полезный текст. Текст необходимо вывести в окно браузера (рис. 1.2.1).

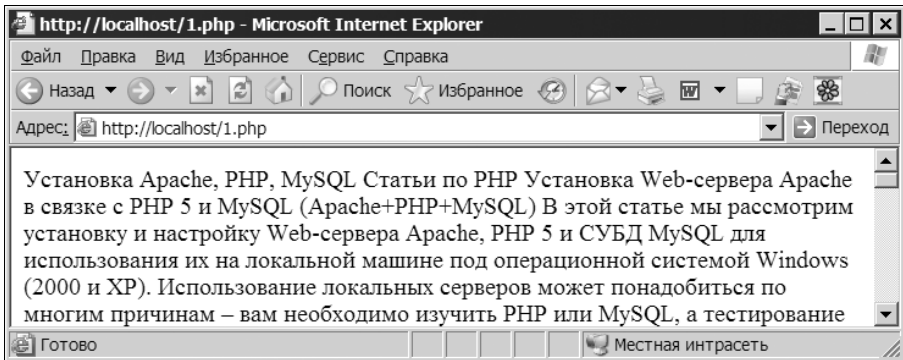


Рис. I.2.1. Чистый текст, извлеченный из HTML-страницы index.htm

## I.2.2. Удаление изображений из HTML-страницы

На компакт-диске найдите HTML-страницу `scripts\2\index.htm`. Прочитайте содержимое страницы и удалите HTML-теги `<img>`.

## I.2.3. Преобразование нескольких пробельных символов в один

Пусть имеется текст, который содержит между словами от одного до нескольких пробельных символов. Необходимо таким образом преобразовать текст, чтобы между словами остался только один пробел.

## I.2.4. Извлечение названия HTML-страницы

На компакт-диске найдите HTML-страницу `scripts\2\index.htm`. Извлеките название страницы, которое помещается между тегами `<title>` и `</title>`.

## I.2.5. Конвертация даты из MySQL-формата в календарный формат

Используя регулярные выражения, переконвертируйте дату из формата `2003-03-21` в формат `21.03.2003`.

## I.2.6. Проверка корректности ввода адреса электронной почты

Разработайте HTML-форму, обработчик которой будет проверять корректность ввода адреса электронной почты (рис. I.2.2).

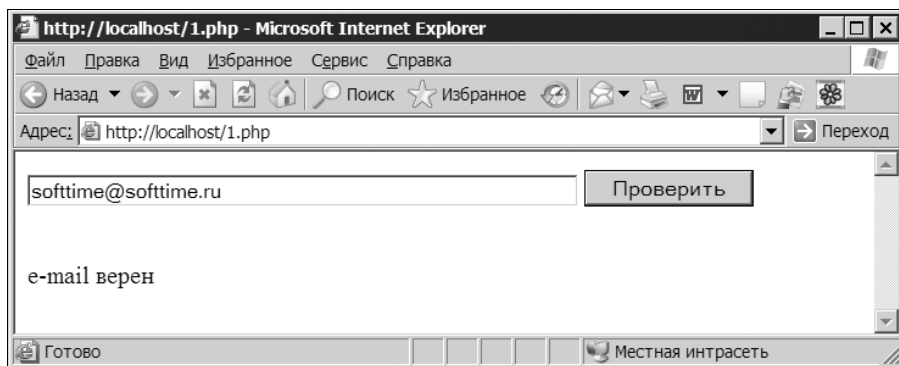


Рис. 1.2.2. HTML-форма проверки адреса электронной почты

## 1.2.7. Проверка корректности ввода URL

Разработайте HTML-форму, обработчик которой будет проверять корректность ввода адреса Web-сайта. Допускается ввод как с указанием протокола, например, **http://www.softtime.ru**, так и без него, например, **www.softtime.ru**. Следует учитывать, что адрес может содержать путь после доменного имени, а также параметры, например, **http://www.softtime.ru/php5/index.php?id\_article=43**.

## 1.2.8. Подсветка URL

Часто возникает задача превращения текстовой ссылки в гиперссылку. На компакт-диске найдите текстовый файл `scripts\2\text.txt` и выведите его содержимое в окно браузера, преобразовав все URL в гиперссылки.

## 1.2.9. Проверка корректности ввода чисел

Создайте HTML-форму, состоящую из двух текстовых полей, в первом из которых вводится количество товарных позиций, а во втором их цена в формате `###.##`. Обработчик формы должен проверить, является ли введенная в первом поле информация целым числом, а во втором — удовлетворяющим денежному формату. Если все верно, необходимо вывести произведение этих двух чисел.

## 1.2.10. Изменение регистра

Пусть имеется фраза "ПРОГРАММИРОВАНИЕ — это ИСКУССТВО. Ему и ЖИЗНЬ посвятить не жалко". Создайте скрипт и регулярное выражение,

которое заменит все слова в верхнем регистре на слова, начинающиеся с заглавной буквы: "Программирование — это Искусство. Ему и Жизнь посвятить не жалко".

### I.2.11. Разбивка длинной строки

При построении различных Web-приложений, главным образом гостевых книг, форумов и чатов, часто возникает необходимость защиты дизайна страниц от длинных последовательностей символов, которые могут исказить дизайн. Создайте функцию, разбивающую на части все слова, длина которых превышает 25 символов.

### I.2.12. Разбивка HTML-страницы на предложения

На компакт-диске найдите HTML-страницу `scripts\2\index.htm`. Прочитайте содержимое страницы и поместите каждое предложение текста в элементы массива `$text` так, чтобы первое предложение оказалось в элементе с индексом 0 — `$text[0]`, второе в элементе с индексом 1 — `$text[1]` и т. д. После чего в цикле преобразуйте массив `$text` в двумерный массив таким образом, чтобы в элементе `$text[0][0]` хранилось первое слово первого предложения, в элементе `$text[0][1]` хранилось второе слово первого предложения и т. д. Проконтролируйте результаты работы, отправив дамп массива в окно браузера при помощи функции `print_r()`.

### I.2.13. Количество слов в тексте

На компакт-диске найдите HTML-страницу `scripts\2\index.htm`. Прочитайте содержимое страницы и сосчитайте, сколько в нем содержится одно-, двух-, ..., десятибуквенных слов.

### I.2.14. Интерпретация тегов bbCode

В Интернете большое распространение получили теги в квадратных скобках, именуемые так же, как теги в стиле phpBB (известного и широко распространенного форума). Удобство использования таких тегов заключается в том, что все теги HTML можно запретить, преобразуя их при помощи функции `htmlspecialchars()` в безопасную форму, и в то же время разрешить посетителям использовать их эквиваленты. Например, `[i]` вместо `<i>` и `[code]` вместо `<code>`. Теги в квадратных скобках можно заменить на теги в угловых скобках уже после преобразования текста при помощи функции



`htmlspecialchars()`. Чаще всего прибегают к тегам `[url]`, которые имеют следующий синтаксис:

```
[url = ссылка] имя ссылки [/url]
```

При выводе на страницу этот шаблон следует преобразовать в

```
<a href=ссылка>имя ссылки</a>
```

Если используется форма тега

```
[url]ссылка[/url]
```

то на страницу выводится гиперссылка вида:

```
<a href=ссылка>ссылка</a>
```

На компакт-диске найдите HTML-страницу `scripts\2\bb.txt`, содержимое этой страницы представлено на рис. 1.2.3.

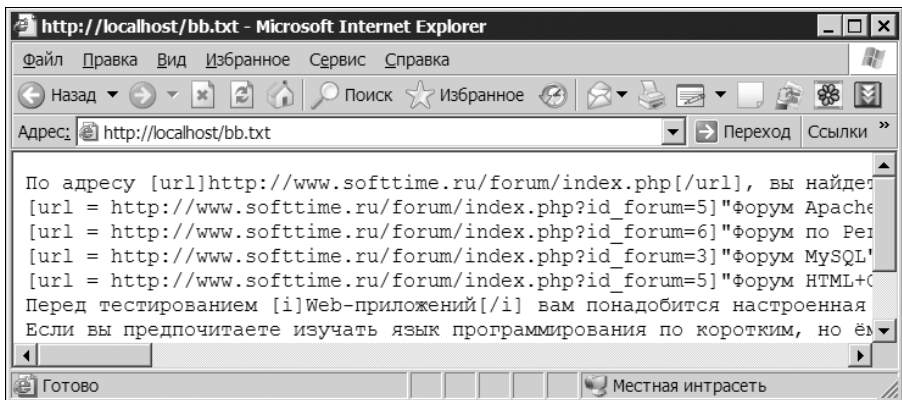


Рис. 1.2.3. Содержимое файла `bb.txt`

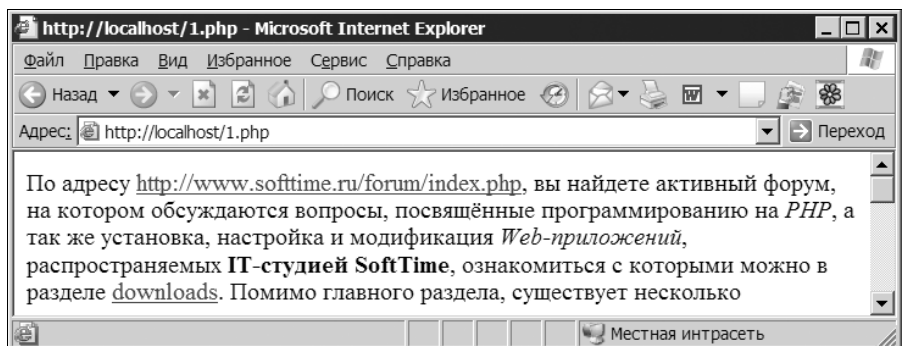


Рис. 1.2.4. Преобразованное содержимое файла `bb.txt`