# R Package Development by Means of Literate Programming (noweb)

Bernhard Pfaff January 17, 2017

# 1 Introduction

The topic of this article is to elucidate the creation of the R package hiker by means of literate programming (henceforth: LP). LP has been introduced by Knuth (1984). Its constituent characteristic is that the source code of a program as well as the documentation thereof resides in the same file. To stay in the example of the **hiker** package: within the file you are currently reading, i.e., hiker.Rnw, the whole R source code of the package as well as an explanation of it, is provided. This file can then be processed by the facilities offered in the R package Rnoweb (see Ihaka, 2013). Rnoweb is a reimplementation of the noweb software (see Ramsey, 1994) written entirely in R. Creating the whole R package hiker by mere processing (tangling of the R code) of this file, is however not entirely true. Given that R packages must meet certain requirements with respect to the directory structure and the documentation of the provided functions, methods and classes, some additional tools become necessary. With respect to setting up the package structure and saving the produced R source files into the package's sub-directory /R, a Makefile is employed (the content of this file is provided in the appendix of this article, see section 4.5). The generation of the man-pages is accomplished by utilizing the facilities provided in the roxygen2 and devtools packages (see Wickham et al., 2015; Wickham and Chang, 2016, respectively).

What is the package **hiker**(pronounce as 'hike—R') about? Within **hiker** the routines for detecting local peaks/troughs of a time series as proposed by Palshikar (2009) have been implemented. The knowledge of these points and the time spans in between these local maxima/minima can be utilized for analyzing the behavior of predictor variables in up-/down markets of a financial time series, *e.g.*, a stock price, for instance. Surely, these phases could in principal be detected by mere eyeball-econometrics, but this approach becomes burdensome if local peaks/troughs of many time series must be determined. The package itself is purely written in R and S4-classes and methods are employed.

In the remainder of this article, the proposed algorithms are presented in the subsequent section. The package structure, the defined S4-classes with the available methods are discussed in Section 3. In the Appendix of this paper the documentation of the package in the roxygen-format is provided. This

<sup>&</sup>lt;sup>1</sup>This file and the requirements for generating the R package **hiker** are made available on github as project lp4rp.

organization has the advantage that the focus is shifted on the R code snippets only and is not interspersed with the documentation of the classes and methods and functions provided by the package.

# 2 Detecting Peaks/Troughs

# 2.1 Notation

A uniformly sampled time series  $\mathbf{y} = \{y_1, \dots, y_i, \dots, y_T\}$  with T data points is considered. The detection of peak/trough points is achieved by a function  $S(i, y_i, T)$  that returns for data point  $y_i$  a score value.<sup>2</sup>. If this score value surpasses a user-provided threshold value  $\theta$ , i.e,  $S(i, y_i, T) \geq \theta$  then the point is considered as a local peak/trough.

Furthermore, in case local peak/trough points appear closely together with respect to time (clustered), then these points can be classified as a burst or bust, respectively.

# 2.2 Algorithms

In Palshikar (2009) five different score functions S have been suggested. All have in common, that a centered window of size 2\*k+1 around  $y_i$  is considered. That is, for a positive integer k the k right neighbors  $N^+(i,k,T) = \{y_{i+1},\ldots,y_{i+k}\}$  and the k left neighbors  $N^-(i,k,T) = \{\{y_{i-k},\ldots,y_{i-1}\}\}$  are employed for assessing  $y_i$  as a local peak/trough. The union of  $N^-(i,k,T)$  and  $N^+(i,k,T)$  is defined as  $N(i,k,T) = N^-(i,k,T) \cdot N^+(i,k,T)$  and if the center point is included as  $N'(i,k,T) = N^-(i,k,T) \cdot y_i \cdot N^+(i,k,T)$ .

The first function,  $S_1$ , computes the score value as the average of the maximum differences between  $y_i$  with its left and right neighbors. The function is defined as:

$$S_1 = \frac{\max(y_i - y_{i-1}, \dots, y_i - y_{i-k}) + \max(y_i - y_{i+1}, \dots, y_i - y_{i+k})}{2}$$
 (1)

For a span of a time series  $\mathbf{y}$  of length 2k+1, denoted by  $\mathbf{x}$ , the equation (1) can be casted in R as:

```
\langle score-maxdiff \ 2 \rangle \equiv \\ scmaxdiff \ \langle -function(x, k) \{ \\ cp \ \langle -k + 1L \\ lmax \ \langle -max(x[cp] - head(x, k)) \\ rmax \ \langle -max(x[cp] - tail(x, k)) \\ (lmax + rmax) \ / \ 2.0 \\ \} \\ Defines: \\ scmaxdiff, used in chunks 5-7.
```

 $<sup>^2</sup>$ It suffices to provide a score function for peaks only. Trough points can be detected by using the negative values of the series y.

Instead of using the maximum differences of  $y_i$  with its k left and right neighbors as in (1), an alternative is to compute the mean differences and evaluate the average thereof:

$$S_2 = \frac{\frac{(y_i - y_{i-1}, \dots, y_i - y_{i-k})}{k} + \frac{(y_i - y_{i+1}, \dots, y_i - y_{i+k})}{k}}{2}$$
(2)

This equation can be casted in R as:

```
3a \langle score\text{-}diffmean \ 3a \rangle \equiv
scdiffmean <- function(x, k){
    cp <- k + 1L
    ldmean <- x[cp] - mean(head(x, k))
    rdmean <- x[cp] - mean(tail(x, k))
    (ldmean + rdmean) / 2.0
}
Defines:
scdiffmean, used in chunks 5-7.
```

Another variation of score computation that has been proposed by Palshikar (2009) is to consider the differences to the mean of the k left and right neighbors, that is:

$$S_3 = \frac{\left(y_i - \frac{(y_{i-1}, \dots, y_{i-k})}{k}\right) + \left(y_i - \frac{(y_{i+1}, \dots, y_{i+k})}{k}\right)}{2} \tag{3}$$

The equation (3) can be casted as R function scavgdiff() for instance as follows:

```
3b \langle score-avgdiff 3b \rangle \equiv
scavgdiff <- function(x, k){

cp <- k + 1L

lmean <- mean(x[cp] - head(x, k))

rmean <- mean(x[cp] - tail(x, k))

(lmean + rmean) / 2.0

}
Defines:
scavgdiff, used in chunks 5-7.
```

The fourth proposed score function differs from the previous three in the sense that it does take the differences between  $y_i$  and its neighbors explicitly into account, but tries to capture its information content by means of relative entropy. The entropy of a vector A with elements  $A = \{a_1, \ldots, a_m\}$  is given as:

$$H_w(A) = \sum_{i=1}^{M} (-p_w(a_i) \log(p_w(a_i)))$$
(4)

where  $p_w(a_i)$  is an estimate of the density value at  $a_i$ . The score function is now based on computing the entropies of H(N((k,i,T))) and H(N'(k,i,T)). Hereby, the densities can be determined by means of a kernel density estimator. The score function is then defined as the difference of the entropies:

$$S_4 = H(N((k, i, T)) - H(N'((k, i, T)))$$
(5)

This concept is implemented in the function scentropy(). The empirical density is computed by calling density(). The ellipsis argument of scentropy() is passed down to this function and hereby allowing the user to employ other than the default arguments of density().

```
 \begin{array}{lll} & \langle score\text{-}entropy \ 4 \rangle \equiv \\ & scentropy \ \leftarrow \ function(x, \ k, \ \ldots) \{ \\ & cp \ \leftarrow \ k + 1L \\ & dfull \ \leftarrow \ density(x, \ \ldots) \$y \\ & hfull \ \leftarrow \ sum(-dfull \ * \ log(dfull)) \\ & dexct \ \leftarrow \ density(x[-cp], \ \ldots) \$y \\ & hexct \ \leftarrow \ sum(-dexct \ * \ log(dexct)) \\ & hfull \ \leftarrow \ hexct \\ & \} \\ & Defines: \\ & scentropy, used in chunks 5-7. \end{array}
```

Finally, a moment-based score function has been put forward in the article by Palshikar. Hereby, the first and second moment of N((k, i, T)) are computed and a t-type statistic can be computed as  $(y_i - m)/s$ . If this statistic surpasses a provided threshold h, then the data point is considered as a local peak/trough.

$$S_5 = \begin{cases} 1 & (y_i - m)/s \ge h \\ 0 & \text{else} \end{cases}$$
 (6)

This type of scoring algorithm is implemented as function scttype() below:

5b

Incidentally, an ensemble forecast of these five algorithms can be utilized for local peak/trough classification. Hereby, one could either use a hybrid approach, whereby only those data points are considered as peak/trough points, if all five methods coincide. This concept is casted in the function schybrid(). Hereby, the signs of all five scoring algorithm are tested for equality.

```
⟨score-hybrid 5b⟩≡
  schybrid <- function(x, k, tval, ...){</pre>
      s <- c(sign(scmaxdiff(x, k)),</pre>
               sign(scavgdiff(x, k)),
               sign(scdiffmean(x, k)),
               sign(scentropy(x, k, ...)),
               sign(scttype(x, k, tval)))
      val <- unique(s)</pre>
      if (length(val) < 2){
           return(s[1])
      } else {
           return(0)
      }
  }
Defines:
  schybrid, used in chunk 7.
Uses scavgdiff 3b, scdiffmean 3a, scentropy 4, scmaxdiff 2, and scttype 5a.
```

It is also conceivable to base the classification on a majority vote. For instance, if three out of the five algorithm classify a data point as a local peak/trough, then this is taken as sufficient evidence. This approach is defined in the function scvote() below. The count of same 'votes' is set by the argument confby. Its default value is 3, *i.e.*, a simple majority. For confby = 5 the function would return the same classification as schybrid() does.

```
\langle score\text{-}vote 6 \rangle \equiv
6
        scvote <- function(x, k, tval, confby = 3, ...){</pre>
             s <- c(sign(scmaxdiff(x, k)),</pre>
                      sign(scavgdiff(x, k)),
                      sign(scdiffmean(x, k)),
                      sign(scentropy(x, k, ...)),
                      sign(scttype(x, k, tval)))
             pos < - rep(1, 5)
             zer \leftarrow rep(0, 5)
             neg < -rep(-1, 5)
             spos \leftarrow sum(s == pos)
             szer <- sum(s == zer)</pre>
             sneg \leftarrow sum(s == neg)
             v <- c(spos, szer, sneg)</pre>
             idx \leftarrow which(v >= confby)
             vals <- c(1, 0, -1)
             if (length(idx) > 0){
                  return(vals[idx])
             } else {
                  return(0)
             }
        }
      Defines:
        scvote, used in chunk 7.
      Uses scavgdiff 3b, scdiffmean 3a, scentropy 4, scmaxdiff 2, and scttype 5a.
```

# 2.3 Combining score methods

What has been accomplished so far, are the function definitions of the five proposed heuristics and combinations thereof with respect to a sub-sample a time series of length 2k+1. These functions are now combined into a single score() function as shown in the subsequent R code chunk. The selection of the kind of score computation is conducted by calling the switch() function.

```
\langle score\text{-}wrapper 7 \rangle \equiv
  score <- function(x, k,</pre>
                       scoreby = c("vote", "avg", "diff", "max", "ent",
                                     "ttype", "hybrid"),
                       tval = 1.0, confby = 3, ...){
      scoreby <- match.arg(scoreby)</pre>
      ans <- switch(scoreby,
                       vote = scvote(x, k, tval, confby, ...),
                       avg = scavgdiff(x, k),
                       diff = scdiffmean(x, k),
                       max = scmaxdiff(x, k),
                       ent = scentropy(x, k, ...),
                       ttype = scttype(x, k, tval),
                       hybrid = schybrid(x, k, tval, ...)
                       )
      ans
  }
Defines:
  score, used in chunk 8b.
Uses scavgdiff 3b, scdiffmean 3a, scentropy 4, schybrid 5b, scmaxdiff 2, scttype 5a,
  and scvote 6.
```

The functions for computing scores are grouped together in the R file score.R as shown in the next chunk. The function definitions are interspersed with the roxygen tags, which will be parsed to the Rd-file score.Rd. Incidentally, only code chunk with names ending in .R are treated as files that should be tangled and saved under the given code chunk's name.

```
\langle score.R \ 8a \rangle \equiv
8a.
               \langle man-func-score 27 \rangle
               \langle score\text{-}wrapper 7 \rangle
              #' Ordname score
               \langle score\text{-}maxdiff 2 \rangle
              #' Ordname score
               \langle score\text{-}diffmean 3a \rangle
              #' Ordname score
               \langle score\text{-}avgdiff 3b \rangle
              #' Ordname score
               \langle score\text{-}entropy 4 \rangle
              #' Ordname score
               \langle score\text{-ttype 5a} \rangle
              #' Ordname score
               \langle score-hybrid 5b \rangle
              #' Ordname score
               \langle score\text{-}vote 6 \rangle
          This code is written to file score.R.
```

So far the wrapper function <code>score()</code> has been created, by which a single point is assessed for being a local maximum or minimum. For analyzing a whole time series for its local extrema, this routine can then be applied to each data point and its left/right neighbors. This task is accomplished with the function <code>hiker()</code> as defined next.

0

The arguments of the function are y for the time series object, k for the count of left/right neighbors, and scoreby for the selection of the scoring method. The arguments tval and confby belong the scoring concepts 'ttype' and 'hybrid', respectively, and the ellipsis argument is passed down to the call of scentropy() for scoreby = 'ent'.

The function body consists of three parts. First, the provided arguments are checked for their validity (as shown in the following code chunk). The computation of the scores is accomplished with the rollapply() function of the package zoo (see Zeileis and Grothendieck, 2005). Finally, the returned object is created.

```
\langle hiker-check 9 \rangle \equiv
      y <- as.zoo(y)
      ## checking arguments
      k <- as.integer(abs(k))</pre>
      ms < -2 * k + 1L
      if ( is.null(dim(y)) ){
          yname <- "series"</pre>
          n <- length(y)</pre>
          if ( n < ms ) {
               stop(paste("Sample size of 'y' is ", n,
                           " and k = ", k, ".\n", sep = ""))
               }
      } else {
          n \leftarrow nrow(y)
          yname <- colnames(y)[1]</pre>
          if ( n < ms ) {
               stop(paste("Sample size of 'y' is ", n,
                           " and k = ", k, ".\n", sep = ""))
          }
          if (ncol(y) > 1) {
               stop("Provide univariate time series of S3-class 'zoo'.\n")
          }
      if ( (confby < 3) || (confby > 5) ){
          stop("\nArgument 'confby' must be integer and in set {3, 4, 5}.\n")
      scoreby <- match.arg(scoreby)</pre>
```

(C)

Within the check section of the function body, the object y is first coerced to a zoo object and the count of neighbors is coerced to a positive integer. Next, the size of the sub-sample for computing the scores is assigned to the object ms. The remaining part consists of checks whether the series is univariate and its length is at least  $2 \times k + 1$ . Finally, the scoring method is determined from the argument scoreby by means of the match.arg() function.

In the final part of the function body of hiker(), the returned object is created as depicted in the subsequent R code chunk.

First, the time series object y is merged with the series of the computed scores and assigned to ans. Next the type of scoring method is stored as character string to des. Finally, an object of S4-class HikeR is created and returned. The definition of this formal class is provided in section 3.2.

Given that hiker() is the core function of the package, it will be stored in its own hiker. R file, as commanded by the next chunk.

```
10b \langle hiker.R \ 10b \rangle \equiv \langle man-func-hiker \ 28a \rangle \langle hiker-func \ 8b \rangle
```

This code is written to file hiker.R.

# 3 Package structure

#### 3.1 Preliminaries

R packages must contain a DESCRIPTION and a NAMESPACE file. Both reside in the package's root directory. As mentioned earlier, only chunks with names ending in '.R' are saved to disk by the tangling function of **Rnoweb**. Therefore, the content of the description file is first, casted in the DESCRIPTION.R file as shown below and is then moved by a directive in the Makefile to the root directory of the package **hiker** without the file's suffix.

```
Package: hiker
Title: Local Peak and Trough of a Time Series
Version: 0.0.0.9000
Authors@R: person("Bernhard", "Pfaff", email = "bernhard@pfaffikus.de",
role = c("aut", "cre"))

Description: Methods for detecting local peaks and troughs of a time series.
Depends: R (>= 3.3.1), zoo, methods
License: GPL-3
Encoding: UTF-8
LazyData: true
This code is written to file DESCRIPTION.R.
```

@

With respect to the creation of the NAMESPACE file, matters are a bit different. The import/export directives contained in this file can be written as roxygen tags and the assembly of the tags and the saving of the file under the package's root directory is conducted by the document() function of devtools. Hence, it suffices to include the tags in a named chunk with suffix '.R'. In the case hiker the import directives are included in the R-file Allclasses.R. As is indicated by the file name, the definitions of the S4-classes will also reside in this file. The skeleton of this this file is provided in the subsequent chunk.

```
11b \langle Allclasses.R \ 11b \rangle \equiv \langle NAMESPACE \ 12a \rangle \langle HikerClass \ 12b \rangle \langle PtbbClass \ 12c \rangle
```

This code is written to file Allclasses.R.

The import directives for the required packages and functions are then given as:

```
#' @import methods
NULL
#' @import zoo
NULL
#' @importFrom graphics plot
NULL
#' @importFrom stats density sd na.omit start end smooth
NULL
#' @importFrom utils head tail
NULL
# Setting old (aka S3) classes
setOldClass("zoo")
```

# 3.2 S4-classes and generics

@

Two S4-classes are defined. The first, HikeR, has been mentioned briefly in the discussion of the core function hiker(). The class consists of four slots, as shown in the subsequent code chunk. Objects of the slot ys are of class zoo and contain the time series and the score value. The slot k is of type integer and holds the count of left/right neighbors. The character slot scoreby provides a literal description of the employed scoring algorithm and the slot yname is a character string of the time series name.

The HikeR-class does not contain explicitly the information whether a certain data point is classified as a peak or trough, or does belong to a certain phase of the time series progression. This lack of explicit information is filled by the S4-class PTBB; it is a short-hand for 'PeaksTroughsBurstBust'. Its slot pt is of informal class zoo and is a logical whether a data point does belong to one the above mentioned points/phases. The slot type is a literal description thereof and the employed threshold value for making this classification is available in the slot h. The class definition is provided in the subsequent code chunk.

```
12c \langle PtbbClass\ 12c \rangle \equiv \langle man\text{-}class\text{-}PTBB\ 29a} \rangle
setClass("PTBB", slots = list(pt = "zoo",
type = "character",
h = "numeric"))
```

Next, the following generic functions are defined in the file Allgenerics.R. The methods of these generics are discussed in the section 3.3.  $\langle Allgenerics.R \ 13 \rangle \equiv$ 13 # generic for extracting peaks

```
setGeneric("peaks", function(object, ...) standardGeneric("peaks"))
# generic for extracting troughs
setGeneric("troughs", function(object, ...) standardGeneric("troughs"))
# generic for extracting bursts
setGeneric("bursts", function(object, ...) standardGeneric("bursts"))
# generic for extracting busts
setGeneric("busts", function(object, ...) standardGeneric("busts"))
# generic for computing ridges
setGeneric("ridges", function(object, ...) standardGeneric("ridges"))
# generic for computing phases
setGeneric("phases", function(object, ...) standardGeneric("phases"))
# generic for extracting topeaks
setGeneric("topeaks", function(object, ...) standardGeneric("topeaks"))
# generic for extracting totroughs
setGeneric("totroughs", function(object, ...) standardGeneric("totroughs"))
# generic for computing runs
setGeneric("runs", function(x) standardGeneric("runs"))
```

#### 3.3 Methods for S4-class 'HikeR'

In this section the S4-methods for objects of class HikeR are discussed. The provided methods are for showing show(), summarizing summary(), retrieval of peaks peaks() and troughs troughs() for this type of objects. Furthermore, the concept of bursts phases (close occurrence of peaks with respect to time) and busts (close occurrence of troughs with respect to time) are defined as methods bursts() and busts(), respectively. Additional methods for characterising the progression of a time series, such as 'ridges', 'phases', 'to-peaks' and 'to-troughs' are provided, too. Finally, a plot()-method is available whereby the user can hightlight/shade the local optima and the phases in between them. All of these methods are contained in the file hiker-methods.R. The skeleton of this file is provided next.

```
14a \langle HikerMethods.R \ 14a \rangle \equiv \langle HikeR-show \ 14b \rangle 
\langle HikeR-summary \ 15a \rangle 
\langle HikeR-peaks \ 15b \rangle 
\langle HikeR-peaks \ 15b \rangle 
\langle HikeR-troughs \ 16a \rangle 
\langle HikeR-bursts \ 16b \rangle 
\langle HikeR-busts \ 17 \rangle 
\langle HikeR-busts \ 18 \rangle 
\langle HikeR-ridges \ 18 \rangle 
\langle HikeR-phases \ 19 \rangle 
\langle HikeR-topeaks \ 20 \rangle 
\langle HikeR-totroughs \ 21 \rangle 
\langle HikeR-plot \ 22 \rangle 
This code is written to file HikerMethods.R.
```

#### 3.3.1 show-method

The show-method provides a brief overview of HikeR-objects. Its definition is given in the following chunk. The roxygen documentation is deferred to the chunk labeled man-HikeR-show and provided in the appendix 4.3.

```
 \begin{array}{lll} & \langle \mathit{HikeR-show} \ 14b \rangle \equiv \\ & \langle \mathit{man-HikeR-show} \ 29b \rangle \\ & & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\ & & & \\
```

a

First, the employed scoring-method and the count of left/right neighbors is printed via the cat() and paste() functions. Second, the summary-method is called, which provides descriptive statistics of the scores and is formally defined in the next section.

#### 3.3.2 summary-method

The summary-method for numerics is applied to the second column of the slot ys, i.e., the computed score values. These numbers are extracted from the zoo object by utilizing the coredata() function. Akin, to the show-method, the within-package documentation is deferred to the appendix.

#### 3.3.3 peaks-method

The peaks-method is the first in a series of methods for classifying the data points in a time series. This and the following methods returns an object of formal class PTBB. Given that a peak is defined as a score value greater than a threshold (the default is h = 0), the core of the method is the logical statement in the first line of the function's body, whether this is true. The remaining slot entries of PTBB are set directly in the to new().

As before, the roxygen documention of this method is listed in the appendix 4.3.

#### 3.3.4 troughs-method

The mirror of peak detection is the trough classification. The method's body is hence defined as the flip-side of the former method, whereby the logical statement is casted now with respect to score values that are less than the threshold value. The default for the threshold is again set to h = 0. The computed logical array is now of type 'trough' and the object of class PTBB is returned in the last line by invoking new().

#### 3.3.5 bursts-method

Bursts are characterized by clusters of local peaks. Whether the sequential occurrence of peaks do constitute a burst-phase, does depend on the maximum allowed distance between two consecutive peak points. This distance is depicted in the burst-method by the argument b with a default value of k, *i.e.* the count of considered left/right neighbors.

```
16b
        \langle HikeR-bursts \ 16b \rangle \equiv
          \langle man\text{-}HikeR\text{-}bursts 29f \rangle
          setMethod("bursts",
               signature(object = "HikeR"),
               function (object, h = 0, b = object@k) {
                    lpts <- object@ys[, 2] > h
                    lidx <- which(lpts == TRUE)</pre>
                    nidx <- length(lidx)</pre>
                    if ( nidx > 1 ){
                         for ( i in 2:nidx ){
                              didx <- lidx[i] - lidx[i - 1]</pre>
                              if ( didx \le b ) {
                                   lpts[(lidx[i]):(lidx[i - 1])] <- TRUE</pre>
                              }
                         }
                    }
                    ans <- zoo(lpts, order.by = index(object@ys))</pre>
                    new("PTBB", pt = ans, type = "burst", h = h)
          })
```

<u>(</u>

Within the definition of the method's body, first the peak points with respect to a threshold h are determined and assigned to the object lpts. Next, the index lidx of these occurrences is created and its length is saved in the object nidx. This information is then used in the following if-clause, to check whether any local peak points are existent in the first place, a necessary condition for the existence of burst phases. Within the for-loop the distance between two consecutive index entries is computed and compared to the bandwidth b. If the difference is less or equal to b, then the intermittent index entries are set to TRUE. Finally, an object of S4-class PTBB is created that holds as a logical the detected burst-phases.

#### 3.3.6 busts-method

A bust episode is characterized by a concentration of trough points with respect to time. As such, it is the flip-side of burst phases. Hence, the skeleton and definition of the bust-method differs only from the burst-method with respect to the detection of local minimum values. Its definition is provided in the subsequent code chunk.

```
\langle HikeR\text{-}busts \ 17 \rangle \equiv
17
         ⟨man-HikeR-busts 30a⟩
        setMethod("busts",
             signature(object = "HikeR"),
             function (object, h = 0, b = object@k) {
                  lpts <- object@ys[, 2] < h</pre>
                  lidx <- which(lpts == TRUE)</pre>
                  nidx <- length(lidx)</pre>
                  if ( nidx > 1 ){
                       for ( i in 2:nidx ){
                            didx <- lidx[i] - lidx[i - 1]
                            if ( didx \le b ) {
                                lpts[(lidx[i]):(lidx[i - 1])] <- TRUE</pre>
                            }
                       }
                  ans <- zoo(lpts, order.by = index(object@ys))</pre>
                  new("PTBB", pt = ans, type = "bust", h = h)
        })
```

### 3.3.7 ridges-method

Data points of a time series can be classified as members of a ridge-like progression, if they exclusively neither belong to a burst or bust phase. This is given in a broader side-way movement. However, it can be possible burst- and bust phases are interlaced with each other (oscillating progression) and this case is also subsumed as a ridge. The ridge-method defined below, employs the previously discussed burst and bust-methods. The two logical arrays of the burst (object burstp) and bust (object burstp) points are then combined in the matrix bbp. The above stated rule is implemented by computing the row sums thereof and checking whether this is different from unity, in which case a data point is classified as being a member of a ridge-phase (object ans). The method does return an object of S4-class PTBB and the slot type is set to 'ridge'.

```
\langle HikeR\text{-}ridges \ 18 \rangle \equiv
18
         ⟨man-HikeR-ridges 30b⟩
        setMethod("ridges",
             signature(object = "HikeR"),
             function (object, h = 0, b = object@k) {
                  N <- nrow(object@ys)
                  k <- object@k
                  bustp <- busts(object, h = h, b = b)@pt</pre>
                  burstp <- bursts(object, h = h, b = b)@pt</pre>
                  bbp <- cbind(bustp, burstp)</pre>
                  ans <- zoo(FALSE, order.by = index(bustp))</pre>
                  ridx <- which ( (rowSums(bbp) > 1) | (rowSums(bbp) < 1) )</pre>
                  ans[ridx] <- TRUE
                  ans[1:object@k] <- NA
                  ans[(N - k + 1):N] \leftarrow NA
                  new("PTBB", pt = ans, type = "ridge", h = h)
        })
```

### 3.3.8 phases-method

The phases-method encompasses the previous three methods and classifies the phases of a time series into the categories 'burst', 'bust' and 'ridge'. The definition of this method is provided in the subsequent code chunk. First, a numeric ans is defined with a length equal to the size of the time series. Next, each of the three methods are called and the detected busts, bursts and ridges are assigned as character strings to the corresponding positions in ans. Finally, the ans numeric is casted as an object of informal class zoo and utilized in the creation of the PTBB S4-class obejct.

```
19
       \langle HikeR\text{-}phases 19 \rangle \equiv
         \langle man\text{-}HikeR\text{-}phases 30c \rangle
         setMethod("phases",
              signature(object = "HikeR"),
              function (object, h = 0, b = object@k) {
                   N <- nrow(object@ys)
                   ans <- rep(NA, N)
                   bustp <- busts(object, h = h, b = b)@pt</pre>
                   ans[which(bustp == TRUE)] <- "bust"</pre>
                   burstp <- bursts(object, h = h, b = b)@pt</pre>
                   ans[which(burstp == TRUE)] <- "burst"</pre>
                   ridgep <- ridges(object, h = h, b = b)@pt
                   ans[which(ridgep == TRUE)] <- "ridge"</pre>
                   ans <- factor(ans)</pre>
                   ans <- zoo(ans, order.by = index(object@ys))</pre>
                   new("PTBB", pt = ans, type = "phase", h = h)
         })
```

### 3.3.9 topeaks-method

```
20
      \langle HikeR\text{-}topeaks \ 20 \rangle \equiv
         \langle man\text{-}HikeR\text{-}topeaks 30d \rangle
         setMethod("topeaks",
             signature(object = "HikeR"),
             function (object, h = 0) {
                  N <- nrow(object@ys)</pre>
                  k <- object@k
                  ans <- zoo(rep(TRUE, N), order.by = index(object@ys))</pre>
                  ans[1:k] <- NA
                  ans[(N - k + 1):N] \leftarrow NA
                  peakp <- peaks(object, h)@pt</pre>
                  pidx <- which(peakp == TRUE)</pre>
                  npidx <- length(pidx)</pre>
                  troup <- troughs(object, h)@pt</pre>
                  tidx <- which(troup == TRUE)</pre>
                  ntidx <- length(tidx)</pre>
                  if ( npidx == 0 ){
                       warning("\nNo local peak points.\n")
                       return(NULL)
                  }
                  if ( ntidx == 0 ){
                       warning("\nNo local trough points.\n")
                       return(NULL)
                  ## if trough comes first, set prior points to FALSE
                  if ( tidx[1] < pidx[1] ){</pre>
                       ans[(k + 1):tidx[1]] <- FALSE
                  for ( i in 1:ntidx ) {
                       previouspeaks <- which(pidx < tidx[i])</pre>
                       countpreviouspeaks <- length(previouspeaks)</pre>
                       if ( countpreviouspeaks > 0 ){
                           maxpos <- which.max(object@ys[pidx[previouspeaks], 1])</pre>
                            ans[(pidx[maxpos] + 1):tidx[i]] <- FALSE</pre>
                           pidx <- pidx[-c(1:countpreviouspeaks)]</pre>
                  new("PTBB", pt = ans, type = "topeak", h = h)
        })
```

@

# 3.3.10 totroughs-method

```
21 ⟨HikeR-totroughs 21⟩≡
⟨man-HikeR-totroughs 30e⟩
setMethod("totroughs",
signature(object = "HikeR"),
function (object, h = 0) {
ans <- topeaks(object, h)
ans@pt <- !ans@pt
ans@type <- "tottrough"
ans
})
```

### 3.3.11 plot-method

```
22
      \langle HikeR\text{-}plot 22\rangle \equiv
        \langle man\text{-}HikeR\text{-}plot 30f \rangle
        setMethod("plot",
            signature(x = "HikeR", y = "missing"),
            function (x, type = c("series", "score", "both", "zoo"),
                       h = 0, b = x@k, phase = c("none", "pt", "bb"),
                       main = NULL, sub = NULL,
                       pt.peak = list(col = "darkgreen", pch = 19, cex = 0.8),
                       pt.trough = list(col = "darkred", pch = 19, cex = 0.8),
                       area.se = list(col = "lightgray"),
                       area.pb = list(col = "seagreen", density = 20),
                       area.tb = list(col = "red2", density = 20),
                       ...){
                 type <- match.arg(type)</pre>
                 phase <- match.arg(phase)</pre>
                 N <- nrow(x@ys)
                 xidx <- 1:N
                 if ( is.null(sub) ){
                     sub <- paste("Score method: ", x@scoreby,</pre>
                                   ", k = ", x@k, sep = "")
                 if ( type == "score" ){
                     srange <- range(na.omit(x@ys[, 2]))</pre>
                     if ( is.null(main) ){
                         main <- paste("Scores of time series:", x@yname)</pre>
                     plot(c(1, N), srange, axes = FALSE,
                          main = main,
                           sub = sub,
                           xlab = "Time", ylab = "Score", type = "n")
                     do.call(graphics::rect, c(list(xleft = 1,
                                                       ybottom = srange[1],
                                                       xright = x@k,
                                                       ytop = srange[2]),
                                                  area.se))
                     do.call(graphics::rect, c(list(xleft = N - x@k + 1,
                                                       ybottom = srange[1],
                                                       xright = N,
                                                       ytop = srange[2]),
                                                  area.se))
                     graphics::lines(coredata(x@ys[, 2]), type = "h", ...)
                     graphics::abline(h = 0)
                     graphics::abline(h = c(h, -h), col = "red")
                     graphics::box()
                     graphics::axis(1, at = xidx, labels = index(x@ys), tick = FALSE)
                     idx <- pretty(srange)</pre>
```

```
graphics::axis(2, at = idx, labels = idx)
if ( type == "series" ){
    yrange <- range(na.omit(x@ys[, 1]))</pre>
    if ( is.null(main) ){
        if ( phase == "none" ){
            main <- paste("Peaks and Troughs of:", x@yname)</pre>
        } else if ( phase == "pt" ){
            main <- paste("To-Peak and To-Trough Phases of:", x@yname)</pre>
        else if ( phase == "bb") {
            main <- paste("Burst and Bust Phases of:", x@yname)</pre>
    }
    plot(c(1, N), yrange, axes = FALSE,
         main = main, sub = sub,
         xlab = "Time", ylab = "", type = "n")
    do.call(graphics::rect, c(list(xleft = 1,
                                     ybottom = yrange[1],
                                     xright = x@k,
                                     ytop = yrange[2]),
                                area.se))
    do.call(graphics::rect, c(list(xleft = N - x@k + 1,
                                     ybottom = yrange[1],
                                     xright = N,
                                     ytop = yrange[2]),
                                area.se))
    graphics::lines(coredata(x@ys[, 1]), ...)
    peakidx <- which(peaks(x, h = h)@pt == TRUE)</pre>
    ynum <- coredata(x@ys[index(x@ys)[peakidx], 1])</pre>
    do.call(graphics::points, c(list(y = ynum, x = peakidx),
                                  pt.peak))
    troughidx <- which(troughs(x, h = h)@pt == TRUE)</pre>
    ynum <- coredata(x@ys[index(x@ys)[troughidx], 1])</pre>
    do.call(graphics::points, c(list(y = ynum, x = troughidx),
                                  pt.trough))
    if ( phase == "bb" ){
        p \leftarrow phases(x, h = h, b = b)
        pchar <- as.character(coredata(p@pt))</pre>
        burstz <- new("PTBB",</pre>
                       pt = zoo(pchar == "burst",
                                 order.by = index(p@pt)),
                       type = "burst",
                       h = p@h
        burstr <- runs(burstz)</pre>
        if (!is.null(burstr)) {
            xleft <- which(as.character(index(burstz@pt)) %in%</pre>
                             as.character(burstr[, "From"]))
            xright <- which(as.character(index(burstz@pt)) %in%</pre>
                             as.character(burstr[, "To"]))
```

}

```
for (i in 1:length(xleft)) {
             do.call(graphics::rect, c(list(xleft = xleft[i],
                                              ybottom = yrange[1],
                                              xright = xright[i],
                                              ytop = yrange[2]),
                                         area.pb))
        }
    }
    bustz <- new("PTBB",</pre>
                  pt = zoo(pchar == "bust",
                           order.by = index(p@pt)),
                   type = "bust",
                  h = p@h
    bustr <- runs(bustz)</pre>
    if (!is.null(bustr)) {
        xleft <- which(as.character(index(bustz@pt)) %in%</pre>
                        as.character(bustr[, "From"]))
        xright <- which(as.character(index(bustz@pt)) %in%</pre>
                         as.character(bustr[, "To"]))
        for (i in 1:length(xleft)) {
             do.call(graphics::rect, c(list(xleft = xleft[i],
                                              ybottom = yrange[1],
                                              xright = xright[i],
                                              ytop = yrange[2]),
                                         area.tb))
        }
    }
}
if ( phase == "pt" ){
    top <- topeaks(x, h)
    topr <- runs(top)</pre>
    tot <- totroughs(x, h)</pre>
    totr <- runs(tot)</pre>
    if ( !is.null(topr) ) {
        xleft <- which(as.character(index(top@pt)) %in%</pre>
                        as.character(topr[, "From"]))
        xright <- which(as.character(index(top@pt)) %in%</pre>
                         as.character(topr[, "To"]))
        for (i in 1:length(xleft)) {
             do.call(graphics::rect, c(list(xleft = xleft[i],
                                              ybottom = yrange[1],
                                              xright = xright[i],
                                              ytop = yrange[2]),
                                         area.pb))
        }
    if (!is.null(totr)) {
        xleft <- which(as.character(index(tot@pt)) %in%</pre>
                        as.character(totr[, "From"]))
        xright <- which(as.character(index(tot@pt)) %in%</pre>
```

```
as.character(totr[, "To"]))
                for (i in 1:length(xleft)) {
                    do.call(graphics::rect, c(list(xleft = xleft[i],
                                                    ybottom = yrange[1],
                                                    xright = xright[i],
                                                    ytop = yrange[2]),
                                               area.tb))
                }
            }
        }
        graphics::box()
        graphics::axis(1, at = xidx, labels = index(x@ys),
                       tick = FALSE)
        idx <- pretty(yrange)</pre>
        graphics::axis(2, at = idx, labels = idx)
    }
    if ( type == "both" ){
        op <- graphics::par(no.readonly = TRUE)</pre>
        graphics::par(mfrow = c(2, 1))
        plot(x, type = "series", h = h, ...)
        plot(x, type = "score", h = h, ...)
        graphics::par(op)
    if ( type == "zoo" ){
        plot(x@ys, ...)
    }
}
```

# 3.4 runs-method for S4-class 'PTBB'

```
\langle PtbbMethods.R \ 26 \rangle \equiv
26
         \langle man-PTBB-runs 31a \rangle
         setMethod("runs",
                     signature(x = "PTBB"), function(x){
                          if ( any(na.omit(x@pt)) ){
                               p <- as.numeric(coredata(x@pt))</pre>
                               run <- rle(p)</pre>
                               peakruns <- which(run$values == 1)</pre>
                               cumidx <- cumsum(run$lengths)</pre>
                               n <- length(cumidx)</pre>
                               xidxright <- cumidx</pre>
                               xidxleft \leftarrow c(1, cumidx + 1)[-c(n + 1)]
                               xidx <- cbind(xidxleft, xidxright)</pre>
                               peakidx <- xidx[peakruns, ]</pre>
                               runscount <- nrow(xidx)</pre>
                               runsidx <- 1:runscount</pre>
                               ans <- sapply(runsidx, function(i)</pre>
                                    x@pt[xidx[i, 1]:xidx[i, 2]])
                               pidx <- which(unlist(lapply(ans, function(r)</pre>
                                   is.element(TRUE, r[1]))))
                               anstrue <- ans[pidx]</pre>
                               per <- lapply(anstrue, function(i)</pre>
                                    data.frame(start(i), end(i)))
                               ans <- data.frame(do.call("rbind", per))</pre>
                               ans[, "Type"] <- x@type</pre>
                               colnames(ans) <- c("From", "To", "Type")</pre>
                               return(ans)
                          } else {
                               return(NULL)
                          }
                     }
         )
```

This code is written to file PtbbMethods.R.

# 4 Appendix

### 4.1 Documentation of functions

```
27
      \langle man\text{-}func\text{-}score \ 27 \rangle \equiv
        #' Basic scoring methods for local minima and maxima
        #'
        #' These are basic functions for evaluating the centre
        #' point of a time series as local minimum or maximum.
        #' Hereby, a score value is computed according to various methods.
        #' If the score is positive, the centre point is tentatively
        #' classified as a local peak.
        #' Incidentally, negative scores indicate a local minima.
        #' @param x \code{numeric}, vector of length \code{2 * k + 1}.
        #' @param k \code{integer}, the count of left/right neighbours.
        #' @param scoreby \code{character}, the scoring method to be used.
        \# ' Oparam tval \code{numeric}, factor for standard deviation band
        #' if \code{scoreby = 'ttype'}.
        #' @param confby \code{integer}, count of minimum vote,
        #' values in the set code{3:5}.
        #' @param ... ellipsis argument.
        #'
        #' @name score
        #' @family scores
        #' @return \code{numeric}, the score value.
        NULL
        #' Ordname score
```

```
28a
       \langle man-func-hiker 28a \rangle \equiv
         #' Peak/trough scores of time series points
         #'
         #' This function computes the score value for each
         #' data point of a time series. The first and last
         #' \code{k} observations are set to \code{NA}.
         #'
         #' @inheritParams score
         #' @param y \code{zoo}, univariate time series.
         \mbox{\tt\#'} @return An object of S4-class \code{HikeR}.
         #' @family scores
         #'
         #' Oreferences Girish K. Palshikar. Simple Algorithms for
         #' Peak Detection in Time-Series. In \emph{Proc. 1st Int. Conf.
         #' Advanced Data Analysis,
         #' Business Analytics and Intelligence}, 2009.
         #'
         #' @examples
         #' TEX <- SP500[, "TEX"]</pre>
         #' ans <- hiker(TEX, k = 8, scoreby = "hybrid", tval = 0.1)</pre>
         #' ans
         #' plot(ans)
         #'
         #' @export
             Documentation of S4-classes
```

```
// wan-class-HikeR 28b)

#' S4 class \code{HikeR}

#'

#' Formal class for classifying local minima and maxima

#' of a time series.

#'

#' @slot ys \code{zoo}, time series with associated scores.

#' @slot k \code{integer}, count of left/right neighbours around centre point.

#' @slot scoreby \code{character}, scoring method.

#' @slot yname \code{character}, name of the series.

#' @exportClass HikeR
```

```
\langle man\text{-}class\text{-}PTBB \text{ 29a} \rangle \equiv
29a
          #' S4 class \code{PTBB}
          #'
          #' Formal class for peaks, troughs, burst, busts and
          #' intermittent phase of a time series.
          #'
          #' @slot pt \code{zoo}, logical: indicating peak/trough points.
          #' @slot type \code{character}, type of point/phase.
          \#' @slot h \code{numeric}, the threshhold for score evaluation.
          #' @exportClass PTBB
        @
               Documentation of S4-methods
        \langle man\text{-}HikeR\text{-}show 29b \rangle \equiv
29b
          #' Ordname HikeR-class
          #' @param object An object of S4 class \code{HikeR}.
          #' @export
        \langle man\text{-}HikeR\text{-}summary 29c \rangle \equiv
29c
          #' Ordname HikeR-class
          #' @aliases summary
          #' @param ... Ellipsis argument.
          #' @export
29d
        \langle man\text{-}HikeR\text{-}peaks 29d \rangle \equiv
          #' Ordname HikeR-class
          #' @aliases peaks
          #' @param h \code{numeric}, the threshold value for scores
          #' to be considered as peaks/troughs.
          #' @return Object of S4-class \code{PTBB}.
          #' @export
        \langle man\text{-}HikeR\text{-}troughs 29e \rangle \equiv
29e
          #' Ordname HikeR-class
          #' @aliases troughs
          #' @export
        @
        \langle man\text{-}HikeR\text{-}bursts 29f \rangle \equiv
29f
          #' Ordname HikeR-class
          #' @aliases bursts
          #' @param b \code{integer}, intermittent count of points between peaks.
          #' @export
```

```
30a
        \langle man\text{-}HikeR\text{-}busts \ 30a \rangle \equiv
          #' Ordname HikeR-class
          #' @aliases busts
          #' @export
        \langle man\text{-}HikeR\text{-}ridges 30b \rangle \equiv
30b
          #' Ordname HikeR-class
          #' @aliases ridges
          #' @export
        \langle man\text{-}HikeR\text{-}phases 30c \rangle \equiv
30c
          #' Ordname HikeR-class
          #' @aliases phases
          #' @export
30d
        \langle man\text{-}HikeR\text{-}topeaks 30d \rangle \equiv
          #' Ordname HikeR-class
          #' @aliases topeaks
          #' @export
        @
        \langle man\text{-}HikeR\text{-}totroughs 30e \rangle \equiv
30e
          #' Ordname HikeR-class
          #' @aliases totroughs
          #' @export
        \langle man\text{-}HikeR\text{-}plot \ 30f \rangle \equiv
30f
          #' Ordname HikeR-class
          #' @aliases plot
          #' @param x An object of S4 class \code{HikeR}.
          #' @param type \code{character}, whether series, scores or both should be plotted.
          #' @param phase \code{character}, whether burst/bust or topeak/totrough phases
          #' should be drawn in series plot.
          #' @param pt.peak \code{list}, named elements are passed to \code{graphics::points()}.
          #' @param main \code{character}, main title of plot.
          #' @param sub \code{character}, sub title of plot
          #' @param pt.trough \code{list}, named elements are passed to \code{graphics::points()}.
          #' @param area.se \code{list}, named elements are passed to \code{graphics::rect()}
          #' for areas of pre- and post sample points.
          #' @param area.pb \code{list}, named elements are passed to \code{graphics::rect()}
          #' for 'to-peak' or 'burst' phases.
          #' @param area.tb \code{list}, named elements are passed to \code{graphics::rect()}
          #' for 'to-trough' or 'bust' phases.
          #' @export
```

```
#' Weekly price data of 476 S&P 500 constituents.
#' The data set was used in the reference below. The authors adjusted
#' the price data for dividends and have removed stocks if two or
#' more consecutive missing values were found. In the remaining cases
#' the NA entries have been replaced by interpolated values.
#'
#'
#' @format A S3-class \code{zoo} object with 265 weekly observations
\mbox{\#'} of 476 members of the S&P 500 index. The sample starts at 2003-03-03
#' and ends in 2008-03-24.
#'
#' @references Cesarone, F. and Scozzari, A. and Tardella, F.: Portfolio
#'
       selection problems in practice: a comparison between linear and
       quadratic optimization models, Working Paper, Universita degli
#'
#'
       Studi Roma Tre, Universita Telematica delle Scienze Umane and
#'
       Universita di Roma, July 2010.
#'
       \url{http://arxiv.org/ftp/arxiv/papers/1105/1105.3594.pdf}
#' @source \url{http://w3.uniroma1.it/Tardella/datasets.html},\cr
#' \url{ http://finance.yahoo.com/}
"SP500"
```

This code is written to file data.R.

#### 4.5 Makefile

```
# Makefile for creating the R package hiker
PKGNAME := hiker
 PKGVERS = \\ (shell sed -n "s/Version: *\\ ([^ ]*\\)/\\ 1/p" ./\\ (PKGNAME)/DESCRIPTION) 
PKGTAR = $(PKGNAME)_$(PKGVERS).tar.gz
{\tt TEXCMD} := pdflatex
RFILES := Allclasses.R Allgenerics.R score.R hiker.R data.R HikerMethods.R PtbbMethods.R
DFILES := SP500.rda
all: deps tex pdf pkg check
tex: $(PKGNAME).tex
pdf: $(PKGNAME).pdf
deps:
        Rscript -e 'if (!require("devtools")) install.packages("devtools")'
        Rscript -e 'if (!require("Rnoweb")) install.packages("Rnoweb_1.1.tar.gz", repos = NULL, type="source")'
$(PKGNAME).tex: $(PKGNAME).Rnw
        Rscript -e 'library(Rnoweb); noweb("$(PKGNAME).Rnw", tangle = FALSE)'
$(PKGNAME).pdf: $(PKGNAME).tex
        $(TEXCMD) $<
        $(TEXCMD) $<
        bibtex $(PKGNAME).aux
        $(TEXCMD) $<
        $(TEXCMD) $<
pkg: $(PKGNAME).Rnw
        Rscript -e 'library(Rnoweb); noweb("$(PKGNAME).Rnw", weave = FALSE)'
# creating package skeleton
        if [ ! -d "$(PKGNAME)" ]; then mkdir $(PKGNAME); fi
        if [ ! -d "(PKGNAME)/R" ]; then mkdir (PKGNAME)/R; fi
        if [ ! -d "$(PKGNAME)/data" ]; then mkdir $(PKGNAME)/data; fi
# handling R files
        find ./\$(PKGNAME)/R/ -type f -delete
        mv DESCRIPTION.R $(PKGNAME)/DESCRIPTION
        mv $(RFILES) $(PKGNAME)/R/
        cp $(DFILES) $(PKGNAME)/data/
# handling man files
        if [ ! -d "$(PKGNAME)/man"]; then mkdir $(PKGNAME)/man; fi
        find ./$(PKGNAME)/man/ -type f -delete
        Rscript -e 'library(devtools); devtools::document(pkg = "./$(PKGNAME)")'
# building the source tarball
        R CMD build $(PKGNAME)
check: pkg
        R CMD check $(PKGTAR)
clean:
        $(RM) -r $(PKGNAME).Rcheck/
        $(RM) $(PKGNAME).aux $(PKGNAME).log $(PKGNAME).out $(PKGNAME).bbl $(PKGNAME).blg
```

# 5 Chunk Index

```
\langle Allclasses.R 11b \rangle
\langle Allgenerics.R \ 13 \rangle
\langle data.R \ 31b \rangle
⟨DESCRIPTION.R 11a⟩
\langle HikeR-bursts 16b\rangle
\langle HikeR\text{-}busts 17 \rangle
\langle hiker-check 9 \rangle
\langle hiker\text{-}func \ 8b \rangle
⟨hiker-output 10a⟩
\langle HikeR\text{-}peaks 15b \rangle
\langle HikeR\text{-}phases 19 \rangle
\langle HikeR\text{-}plot\ 22\rangle
\langle HikeR\text{-}ridges \ 18 \rangle
\langle HikeR\text{-}show 14b \rangle
\langle HikeR-summary 15a\rangle
\langle HikeR-topeaks 20\rangle
\langle HikeR-totroughs 21\rangle
⟨HikeR-troughs 16a⟩
\langle hiker.R \ 10b \rangle
\langle HikerClass \ 12b \rangle
⟨HikerMethods.R 14a⟩
\langle man\text{-}class\text{-}HikeR 28b \rangle
\langle man\text{-}class\text{-}PTBB 29a \rangle
⟨man-func-hiker 28a⟩
\langle man-func-score 27 \rangle
\langle man\text{-}HikeR\text{-}bursts 29f \rangle
⟨man-HikeR-busts 30a⟩
\langle man\text{-}HikeR\text{-}peaks 29d \rangle
\langle man\text{-}HikeR\text{-}phases 30c \rangle
\langle man\text{-}HikeR\text{-}plot 30f \rangle
\langle man\text{-}HikeR\text{-}ridges 30b \rangle
\langle man\text{-}HikeR\text{-}show 29b \rangle
\langle man\text{-}HikeR\text{-}summary 29c \rangle
\langle man\text{-}HikeR\text{-}topeaks 30d \rangle
\langle man\text{-}HikeR\text{-}totroughs 30e \rangle
\langle man\text{-}HikeR\text{-}troughs 29e \rangle
\langle man-PTBB-runs 31a \rangle
(NAMESPACE 12a)
\langle PtbbClass\ 12c \rangle
\langle PtbbMethods.R \ 26 \rangle
\langle score\text{-}avgdiff 3b \rangle
\langle score\text{-}diffmean 3a \rangle
\langle score\text{-}entropy \ 4 \rangle
\langle score-hybrid 5b \rangle
\langle score\text{-}maxdiff 2 \rangle
\langle score\text{-}ttype 5a \rangle
⟨score-vote 6⟩
```

 $\langle score\text{-}wrapper 7 \rangle$  $\langle score.R 8a \rangle$ 

#### Identifier Index 6

 $\mathtt{scavgdiff:} \ \underline{3b},\, 5b,\, 6,\, 7$  $\mathtt{scdiffmean:} \ \ \underline{3a},\, 5b,\, 6,\, 7$ scentropy:  $\underline{4}$ , 5b, 6, 7 schybrid:  $\underline{5b}$ , 7 scmaxdiff:  $\underline{2}$ , 5b, 6, 7 score:  $\underline{7}$ , 8bscttype:  $\underline{5a}$ , 5b, 6, 7 scvote:  $\underline{6}$ , 7

# References

- Ihaka, R. (2013). Rnoweb: Simplified Literate Programming. R package version 1.1.
- Knuth, D. E. (1984). Literate programming. The Computer Journal 27(2), 97–111.
- Palshikar, G. (2009). Simple algorithms for peak detection in time-series. In First Int. Conf. Advanced Data Analysis, Business Analytics and Intelligence, Ahmedabad, India.
- Ramsey, N. (1994, September). Literate programming simplified. *IEEE Software* 11(5), 97–105.
- Wickham, H. and W. Chang (2016). devtools: Tools to Make Developing R Packages Easier. R package version 1.12.0.
- Wickham, H., P. Danenberg, and M. Eugster (2015). roxygen2: In-Source Documentation for R. R package version 5.0.1.
- Zeileis, A. and G. Grothendieck (2005). zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software* 14(6), 1–27.