

Package ‘rpart.plot’

September 26, 2016

Version 2.1.0

Title Plot 'rpart' Models: An Enhanced Version of 'plot.rpart'

Author Stephen Milborrow

Maintainer Stephen Milborrow <milbo@sonic.net>

Depends rpart (>= 4.1-0)

Suggests earth

Description Plot 'rpart' models. Extends plot.rpart() and text.rpart() in the 'rpart' package.

License GPL-3

URL <http://www.milbo.org/rpart-plot>

NeedsCompilation no

Repository CRAN

Date/Publication 2016-09-26 21:26:49

R topics documented:

prp	2
ptitanic	14
rpart.plot	16
rpart.plot.version1	20
show.prp.palettes	21
Index	22

prp

*Plot an rpart model.***Description**

Plot an `rpart` model.

This function combines and extends `plot.rpart` and `text.rpart` in the `rpart` package. For an overview, please see the package vignette [Plotting rpart trees with the rpart.plot package](#).

The function `rpart.plot` provides a simplified interface to this function. First-time users should use `rpart.plot` rather than this function.

The arguments of this function are a superset of those of `rpart.plot` and some of the arguments have different defaults. See the `rpart.plot` help page for a table showing the different defaults.

(The defaults are different for historical reasons: for backwards compatibility the defaults of `prp` haven't changed, whereas the defaults of `rpart.plot` were changed when `type="auto"` and `box.palette` were introduced in version 2.0.0 of this package.)

Usage

```
prp(x=stop("no 'x' arg"),
    type=0, extra=0, under=FALSE, clip.right.labs=TRUE,
    nn=FALSE, ni=FALSE, yesno=TRUE,
    fallen.leaves=FALSE, branch=if(fallen.leaves) 1 else .2,
    uniform=TRUE, left=TRUE, xflip=FALSE, yflip=FALSE,
    Margin=0, space=1, gap=NULL,
    digits=2, varlen=-8, faclen=3,
    cex=NULL, tweak=1,
    compress=TRUE, ycompress=uniform,
    trace=FALSE, snip=FALSE, snip.fun=NULL,

    box.col=0, box.palette=0,
    pal.thresh=NULL, pal.node.fun=FALSE,
    border.col=col,
    round=NULL, leaf.round=NULL,
    shadow.col=0, prefix="", suffix="", xsep=NULL,

    under.font=font, under.col=1, under.cex=.8,

    split.cex=1, split.font=2, split.family=family, split.col=1,
    split.box.col=0, split.border.col=0,
    split.lty=1, split.lwd=NULL, split.round=0,
    split.shadow.col=0,
    split.prefix="", right.split.prefix=NULL,
    split.suffix="", right.split.suffix=NULL,
    facsep=",", eq=" = ", lt=" < ", ge=" >= ",
```

```

branch.col=if(identical(branch.type, 0)) 1 else "gray",
branch.lty=1, branch.lwd=NULL,
branch.type=0, branch.tweak=1,
min.branch.width=.002, branch.fill=branch.col,

nn.cex=NULL, nn.font=3, nn.family="", nn.col=1,
nn.box.col=0, nn.border.col=nn.col,
nn.lty=1, nn.lwd=NULL, nn.round=.3,
yes.text="yes", no.text="no",

node.fun=NULL,
split.fun=NULL,
FUN="text",

nspace=branch, minbranch=.3, do.par=TRUE,
add.labs=TRUE, clip.left.labs=FALSE, fam.main="",
yshift=0, yspace=space, shadow.offset=.4,
split.adj=NULL, split.yshift=0, split.space=space,
split.yspace=yspace, split.shadow.offset=shadow.offset,
nn.adj=.5, nn.yshift=0, nn.space=.8, nn.yspace=.5,
ygap=gap/2, under.ygap=.5, yesno.yshift=0,
xcompact=TRUE, ycompact=uniform, xcompact.ratio=.8, min.inter.height=4,
max.auto.cex=1, min.auto.cex=.15, ycompress.cex=.7, accept.cex=1.1,
shift.amounts=c(1.5, 2),
Fallen.yspace=.1, boxes.include.gap=FALSE,
legend.x=NULL, legend.y=NULL, legend.cex=1,
...)

```

Arguments

x	An <code>rpart</code> object. The only required argument.
type	Type of plot. Possible values: <ul style="list-style-type: none"> 0 Default. Draw a split label at each split and a node label at each leaf. 1 Label all nodes, not just leaves. Similar to <code>text.rpart</code>'s <code>all=TRUE</code>. 2 Like 1 but draw the split labels below the node labels. Similar to the plots in the CART book. 3 Draw separate split labels for the left and right directions. 4 Like 3 but label all nodes, not just leaves. Similar to <code>text.rpart</code>'s <code>fancy=TRUE</code>. See also <code>clip.right.labs</code>.
extra	Display extra information at the nodes. Possible values: <ul style="list-style-type: none"> "auto" (case insensitive) Automatically select a value based on the model type, as follows: <ul style="list-style-type: none"> <code>extra=106</code> class model with a binary response <code>extra=104</code> class model with a response having more than two levels <code>extra=100</code> other models

- 0** Default. No extra information.
- 1** Display the number of observations that fall in the node (per class for class objects; prefixed by the number of events for poisson and exp models). Similar to `text.rpart`'s use `.n=TRUE`.
- 2** Class models: display the classification rate at the node, expressed as the number of correct classifications and the number of observations in the node. Poisson and exp models: display the number of events.
- 3** Class models: misclassification rate at the node, expressed as the number of incorrect classifications and the number of observations in the node.
- 4** Class models: probability per class of observations in the node (conditioned on the node, sum across a node is 1).
- 5** Class models: like 4 but do not display the fitted class.
- 6** Class models: the probability of the second class only. Useful for binary responses.
- 7** Class models: like 6 but do not display the fitted class.
- 8** Class models: the probability of the fitted class.
- 9** Class models: the probabilities times the fraction of observations in the node (the probability relative to all observations, sum across all leaves is 1).
- +100** Add 100 to any of the above to also display the percentage of observations in the node. For example `extra=101` displays the number and percentage of observations in the node. Actually, it's a weighted percentage using the weights passed to `rpart`.

Note 1: Unlike `text.rpart`, by default `prp` uses its own routine for generating node labels (not the function attached to the object). See the `node.fun` argument.

Note 2: The `extra` argument has special meaning for `mvpart` objects. See the Appendix to this package's vignette.

<code>under</code>	Applies only if <code>extra > 0</code> . Default FALSE, meaning put the extra text <i>in</i> the box. Use TRUE to put the text <i>under</i> the box. See also <code>under.cex</code> .
<code>clip.right.labs</code>	Default is TRUE meaning "clip" the right-hand split labels, i.e. do not print <code>variable=</code> . Applies only if <code>type=3</code> or <code>4</code> . See also <code>clip.left.labs</code> .
<code>nn</code>	Display the node numbers. Default FALSE. (In the current implementation some overplotting may occur with <code>nn=TRUE</code> .)
<code>ni</code>	Display the node indices, i.e. the row numbers of the nodes in the object's frame . Default FALSE.
<code>yesno</code>	One of <ul style="list-style-type: none"> 0 don't write yes and no on the tree 1 (default) write yes and no at the top split 2 write yes and no at all splits. Seems to work best with <code>fallen.leaves=TRUE</code>, <code>split.border.col=1</code>. (The <code>yesno</code> argument is ignored if <code>type=3</code> or <code>4</code>. Use <code>nn.col</code> and the other <code>nn</code> parameters to change the color etc. of the yes and no text. Use <code>yes.text</code> and <code>no.text</code> to change the actual text displayed.)
<code>fallen.leaves</code>	Default FALSE. If TRUE, position the leaf nodes at the bottom of the graph.

branch	Controls the shape of the branch lines. Specify a value between 0 (V shaped branches) and 1 (square shouldered branches). Default is <code>if(fallen.leaves) 1 else .2</code> .
uniform	If TRUE (the default), the vertical spacing of the nodes is uniform. If FALSE, the nodes are spaced proportionally to the fit (more precisely, to the difference between a node's deviance and the sum of its two children's deviances). Very small vertical spaces are automatically artificially expanded to make room for the labels, see <code>minbranch</code> . Note: <code>uniform=FALSE</code> with <code>cex=NULL</code> (the default) can sometimes cause very small text.
left	Default TRUE, meaning the left side of a split is the path taken if the split condition is true. With <code>left=FALSE</code> the split labels are changed so the right side is true.
xflip	Default FALSE. If TRUE, flip the tree horizontally.
yflip	Default FALSE. If TRUE, flip the tree vertically, so the root is at the bottom.
Margin	Extra white space around the tree, as a fraction of the graph width. Default 0, meaning no extra space. To add say 10% space around the tree use <code>Margin=0.1</code> . (This is the <code>margin</code> argument of <code>plot.rpart</code> . The name was changed to prevent partial matching with <code>mar</code> , which can be passed in as a ... argument.)
gap	Minimum horizontal gap between the (possibly invisible) boxes, in character widths. Default NULL, meaning automatically choose a suitable value (normally 1, but if the graph is very crowded will be set to 0, permitting boxes to touch to allow a bigger <code>cex</code>). See also <code>space</code> .
digits	The number of significant digits in displayed numbers. Default 2. If 0, use <code>getOption("digits")</code> . Details: Numbers from 0.001 to 9999 are printed without an exponent (and the number of digits is actually only a suggestion, see <code>format</code> for details). Numbers out that range are printed with an "engineering" exponent (a multiple of 3).
varlen	Length of variable names in text at the splits (and, for class responses, the class in the node label). Default -8, meaning truncate to eight characters. Possible values: 0 use full names. greater than 0 call <code>abbreviate</code> with the given <code>varlen</code> . less than 0 truncate variable names to the shortest length where they are still unique, but never truncate to shorter than <code>abs(varlen)</code> .
faclen	Length of factor level names in splits. Default 3, meaning <code>abbreviate</code> to three characters. Possible values are as <code>varlen</code> above, except that for back-compatibility with <code>text.rpart</code> the special value 1 means represent the factor levels with alphabetic characters (a for the first level, b for the second, etc.).
cex	Default NULL, meaning calculate the text size automatically.
tweak	Adjust the (possibly automatically calculated) <code>cex</code> . Default 1, meaning no adjustment. Use say <code>tweak=1.2</code> to make the text 20% larger. However, since font sizes are discrete the <code>cex</code> you ask for may not be exactly the <code>cex</code> you get. And a

small change to tweak may not actually change the type size, or change it more than you want.

compress	If TRUE (the default), make more space by shifting nodes horizontally where space is available. This often allows larger text. (This is the same as <code>plot.rpart</code> 's argument of the same name, except that here the default is TRUE.)
ycompress	If TRUE (the default unless <code>uniform=FALSE</code>), make more space by shifting labels vertically where space is available. Actually, this only kicks in if the initial automatically calculated <code>cex</code> is less than 0.7. Use <code>ycompress=FALSE</code> if you feel the resulting display is too messy. In the current implementation, the shifting algorithm works a little better (allowing larger text) with <code>type=1, 2, or 3</code> .
trace	Default FALSE. Use TRUE to print the automatically calculated <code>cex</code> , <code>xlim</code> , and <code>ylim</code> . Use integer values greater than 1 for more detailed tracing.
snip	Default FALSE. Set TRUE to interactively trim the tree with the mouse. See the package vignette (or just try it).
snip.fun	Function invoked after each mouse click when <code>snip=TRUE</code> . Default NULL, meaning no function. Otherwise set <code>snip.fun</code> to your own function with the prototype <code>function(tree)</code> , where <code>tree</code> is the snipped tree. See the package vignette for an example.
The following control the node labels.	
box.col	Color of the boxes around the text. Default 0, meaning use the background color. If this argument is used, the <code>box.palette</code> argument is ignored.
box.palette	<p>Palette for coloring the node boxes based on the fitted value. This is a vector of colors, for example <code>box.palette=c("green", "green2", "green4")</code>. Small fitted values are displayed with colors at the start of the vector; large values with colors at the end.</p> <p>The special value <code>box.palette=0</code> (default for <code>prp</code>) uses the background color (typically white).</p> <p>The special value <code>box.palette="auto"</code> (default for <code>rpart.plot</code>, case insensitive) automatically selects a predefined palette based on the type of model.</p> <p>Otherwise specify a predefined palette e.g. <code>box.palette="Grays"</code> for the predefined gray palette (a range of grays). The predefined palettes are (see the show.prp.palettes function):</p> <p>Grays Greys Greens Blues Browns Oranges Reds Purples Gy Gn Bu Bn Or Rd Pu (alternative names for the above palettes) BuGn GnRd BuOr etc. (two-color diverging palettes: any combination of two of the above palettes) RdYlGn GnYlRd (three color palettes)</p> <p>Prefix the palette name with "-" to reverse the order of the colors e.g. <code>box.palette="-auto"</code> or <code>box.palette="-Grays"</code>.</p> <p>The <code>box.palette</code> argument is ignored if the <code>box.col</code> palette argument is specified.</p>
pal.thresh	Applies when <code>box.palette</code> is a two-color diverging palette (such as <code>BuGn</code>). Specifies the response threshold to split the two sub-palettes (such as <code>Bu</code> and <code>Gn</code>).

For example, to display fitted values less than 90 in shades of blue and values greater than 90 in shades of green, use `pal.thresh=90` with `box.palette="BuGn"`. By default `pal.thresh` is calculated automatically. (For a two-class response the default threshold is 0.5; for a continuous response the default is the median fitted value.)

Node boxes for fitted values less than the threshold are displayed using colors from the first sub-palette; boxes for fitted values greater than the threshold are displayed using colors from the second sub-palette.

This argument is ignored if `box.palette` isn't a two-color diverging palette, and is ignored for models with multiple-class responses (more than two classes).

<code>pal.node.fun</code>	Specifies how the <code>box.palette</code> argument is handled when the <code>node.fun</code> argument is specified. Default is FALSE, meaning ignore <code>node.fun</code> and use the fitted value to select the node color from <code>box.palette</code> as usual. If TRUE, use the label returned by the node function (instead of the fitted value) to select the color. The first number in each label returned by the node function is used, skipping over any non-numeric initial text in the label. An error message will be issued if a label doesn't include a number. This argument is ignored if <code>node.fun</code> isn't specified, and is ignored for models with multiple-class responses (more than two classes).
<code>border.col</code>	Color of the box border around the text. Default <code>col</code> , the color of the text in the box. Use 0 for no border. (Note: <code>par</code> settings like <code>col</code> can be passed in as ... arguments. If not passed in, <code>par("col")</code> is used.)
<code>round</code>	Controls the rounding of the corners of the node boxes. Default NULL, meaning calculate automatically. Else specify 0 for sharp edges, and values greater than 0 for rounded edges. Bigger is more round. Values too big for the size of the box get silently reduced.
<code>leaf.round</code>	Controls the rounding of the corners of the leaf node boxes. Default NULL, meaning use <code>round</code> . Else specify a value greater than or equal to 0.
<code>shadow.col</code>	Color of the shadow under the boxes. Default 0, no shadow. Try "gray" or "darkgray". (Note: overlapping shadows look better on devices that support alpha channels. If you get the message "Warning: semi-transparency is not supported" please let me know – it means that a fix is needed to the code that determines if the device supports alpha channels.)
<code>prefix</code>	Default "". Prepend this string to the node labels. So could be the name of the fitted response, for instance.
<code>suffix</code>	Default "". Append this string to the node labels. Text after a double newline "\n\n" (if any) will be plotted <i>under</i> the box. (Actually, double newlines can be used in any of the prefix or suffix arguments for this purpose.)
<code>xsep</code>	String which separates the individual counts and probabilities in node labels when <code>extra>0</code> . Default NULL meaning automatically select: usually " " (two spaces), but " / " for rates. Use <code>xsep="/"</code> for compatibility with <code>text.rpart</code> . See also <code>facsep</code> , which separates the factor levels in split labels.

The following control the text under the boxes (apply only if `under=TRUE` or there is a double newline `\n\n` in `prefix` or `suffix`).

<code>under.font</code>	Font of the text under the box. Default font (which can be passed in as a ... argument).
<code>under.col</code>	Color of the text under the box. Default 1.
<code>under.cex</code>	Size of the text under the box relative to the text in the box. Default .8, smaller than the text in the box.

The following control the split labels.

<code>split.cex</code>	Size of the split text relative to <code>cex</code> (which by default is calculated automatically). Default 1.
<code>split.font</code>	Font for the split labels. Default 2, bold. (Note: use <code>font</code> to change the <i>node</i> label text.)
<code>split.family</code>	Font family for the split labels. Default "", or use something like <code>split.family="serif"</code> . (Note: use <code>family</code> to change the <i>node</i> label text.)
<code>split.col</code>	Color of the split label text. Default 1. (Note: use <code>col</code> to change the <i>node</i> label text.)
<code>split.box.col</code>	Color of the split boxes. Default 0, meaning use the background color.
<code>split.border.col</code>	Color of the split box borders. Default 0, invisible.
<code>split.lty</code>	Line type for the split box borders. The default is 1, but the border will be invisible unless you change the default <code>split.border.col</code> . (Note: use <code>lty</code> to change the <i>node</i> box borders.)
<code>split.lwd</code>	Line width of the split box border relative to <code>cex</code> (which by default is calculated automatically). The border is by default invisible, see <code>codesplit.border.col</code> .
<code>split.round</code>	Controls the rounding of the corners of the split boxes. Default 0, meaning sharp corners. Else specify a value greater than or equal to 0. The split boxes are by default invisible, see <code>split.box.col</code> and <code>split.border.col</code> .
<code>split.shadow.col</code>	Color of the shadow under the split boxes. Default 0, no shadow.
<code>split.prefix</code>	Default "". Prepend this string to the split labels.
<code>right.split.prefix</code>	Default <code>split.prefix</code> . Prepend this string to the right split labels. Applies only when <code>type=3</code> or <code>4</code> .
<code>split.suffix</code>	Default "". Append this string to the split labels.
<code>right.split.suffix</code>	Default <code>split.suffix</code> . Append this string to the right split labels. Applies only when <code>type=3</code> or <code>4</code> .
<code>facsep</code>	Default ", ". String which separates the factor levels in split labels. See also <code>xsep</code> , which separates the individual counts when <code>extra</code> is used.
<code>eq</code>	Default " = ". String which separates the factor name from the levels in split labels. The idea is that you can add or remove spaces around the =, or use words if that suits you.
<code>lt</code>	Default " < ". String which represents "less than" in split labels.

ge Default " >= ". String which represents "greater than or equal" in split labels.

The following control the branches.

branch.col Color of the branch lines. Default 1, but set to "gray" if branch.type is nonzero.

branch.lty Branch line type. Default 1.

branch.lwd Line width of the branch lines relative to cex (which by default is calculated automatically). (Note: branch.lwd does not control the width of the "wide branches" drawn when branch.type is nonzero.)

branch.type Default 0. If nonzero draw "wide branches", with branch widths proportional to the parameter selected by branch.type as follows:

0 The default. The branch lines are drawn conventionally.

1 deviance

2 sqrt(deviance)

3 deviance / nobs

4 sqrt(deviance / nobs) (the standard deviation when method="anova")

5 weight (frame\$wt). This is the number of observations at the node, unless rpart's weight argument was used.

6 complexity

7 abs(predicted value)

8 predicted value - min(predicted value)

9 constant (for checking visual perception of the relative width of branches).

Otherwise set branch.type to your own function. The function should take a single argument x (the rpart object) and return a numeric vector of non-negative widths corresponding to rows in frame. See get.branch.widths in the source code.

Note: with a nonzero branch.type, in the current implementation the branch argument will be silently changed to 1 (if branch > .5) or 0 (if branch < .5)

branch.tweak Default 1. Applies only if branch.type is nonzero. Use this argument to scale the widths of the branches, for example, branch.tweak=.5 to halve the width of the branches. (By default, prp normalizes the widths so the widest branch is one-fifth the plot width.)

min.branch.width Default 0.002. Applies only if branch.type is nonzero. The minimum width of a branch, as a fraction of the page width. The width of branches that would be thinner than min.branch.width is clamped. Increase min.branch.width if the thinnest branches are too skinny on your display device.

branch.fill Color used to fill the wide branch lines. Applies only if branch.type is nonzero. Default branch.col.

The following control the node numbers (with nn=TRUE).

nn.cex Default NULL, meaning calculate the cex of the node numbers automatically. This and the following arguments apply only when nn=TRUE.

<code>nn.font</code>	Font for the node numbers. Default 3, italic.
<code>nn.family</code>	Font family for the node numbers. Default "".
<code>nn.col</code>	Color of the node number text. Default 1.
<code>nn.box.col</code>	Color of the boxes around the node numbers. Default 0, meaning use the background color.
<code>nn.border.col</code>	Color of the box border around the node numbers. Default <code>nn.col</code> .
<code>nn.lty</code>	Line type of the node number box border. Default 1.
<code>nn.lwd</code>	Line width of the node box border relative to <code>cex</code> (which by default is calculated automatically). Default NULL, meaning use <code>lwd</code> (which can be passed in as a ... argument).
<code>nn.round</code>	Controls the rounding of the corners of the node number boxes. Default .3, meaning small corners. Else specify a value greater than or equal to 0.
<code>yes.text</code> , <code>no.text</code>	Text displayed when <code>yesno=TRUE</code> . Default <code>yes.text="yes"</code> and <code>no.text="no"</code> .
	The following are user definable functions.
<code>node.fun</code>	The function that generates the text at the node labels. The default is NULL, which means use a default function internal to prp. (This is necessary for full support of <code>extra</code> as described in the section on <code>extra</code> above.) Otherwise set <code>node.fun</code> to your own function with the prototype <code>function(x, labs, digits, varlen)</code> See the package vignette for details. See also the <code>pal.node.fun</code> argument.
<code>split.fun</code>	The function that generates the text at the splits. The default is NULL, which means use a default function internal to prp. Otherwise set <code>split.fun</code> to your own function with the prototype <code>function(x, labs, digits, varlen, faclen)</code>
<code>FUN</code>	The function that displays the text on the screen. Default text .
	The following are esoteric parameters, mostly for the graph layout engine.
<code>nspace</code>	Applies only when <code>compress=TRUE</code> . Default <code>nspace=branch</code> . The size of the space between a split and a leaf, relative to the space between leaves.
<code>minbranch</code>	Applies only when <code>uniform=FALSE</code> . Default .3. The minimum height between levels is clamped at <code>minbranch</code> times the mean interlevel distance. Needed because sometimes a split gives little or no improvement in deviance, and an interlevel distance strictly proportional to the improvement would leave no room for the label.
<code>do.par</code>	Default TRUE, meaning adjust the <code>mar</code> parameter so the tree fills the figure region. This also sets <code>xpd=NA</code> . These graphic parameters are restored to their original state before prp exits. If you explicitly set <code>mar</code> or <code>xpd</code> , prp will use your setting regardless of the setting of <code>do.par</code> .
<code>add.labs</code>	Default TRUE, meaning display the labels. If FALSE, gives a bare bones display similar to <code>plot.rpart</code> .
<code>clip.left.labs</code>	Like <code>clip.right.labs</code> but for the left labels. Default is FALSE. Note that <code>clip.left.labs</code> and <code>clip.right.labs</code> can be vectors, indexed on the split number.

<code>fam.main</code>	Font family for the main text. Default <code>""</code> . The (inconsistent) name was chosen to minimize partial matching with <code>main</code> and <code>family</code> which can be passed in as in as ... arguments.
<code>yshift</code>	Vertical position of the labels, in character heights relative to their default position. Default <code>0</code> . Negative values move the text down; positive up (the box around the text will follow along).
<code>space</code>	Horizontal space to the box border on each side of the node label text, in character widths. Default <code>1</code> . Use this (and <code>yspace</code>) for bigger boxes. Since this affects the size of the (possibly invisible) boxes, it also affects the graph layout and hence also the automatic calculation of <code>cex</code> .
<code>yspace</code>	Vertical space to the box border above and below the node label text, in character heights. Default <code>space</code> . See the comments for <code>space</code> .
<code>shadow.offset</code>	Offset of the shadow from the boxes, in character widths. Default <code>.4</code> (but the shadow will be invisible unless the default <code>shadow.col</code> is changed).
<code>split.adj</code>	Horizontal position of the split text. In string width units, as is the convention for <code>adj</code> arguments. Default <code>NULL</code> , meaning use <code>adj</code> (which defaults to <code>0.5</code> but can be passed in as a ... argument). Use values less/more than <code>.5</code> to shift the text left/right (the box around the text will follow along).
<code>split.yshift</code>	Vertical position of the split labels, in character heights relative to their default positions. Default <code>0</code> . Negative values move the text down; positive up (the box around the text will follow along). This adjusts the positions of the split labels relative to the node labels. (Use <code>yshift</code> if you want to shift <i>both</i> the split and node labels.)
<code>split.space</code>	Horizontal space between the split label text and the box, in character widths. Default <code>space</code> . Affects the size of the box drawn around the text. The split boxes are by default invisible (see <code>split.box.col</code> and <code>split.border.col</code>), but nevertheless affect the graph layout used in the automatic calculation of <code>cex</code> .
<code>split.yspace</code>	Vertical space between the split label text and the box, in character heights. Default <code>yspace</code> .
<code>split.shadow.offset</code>	Offset of the shadow from the split boxes, in character widths. Default <code>shadow.offset</code> . (but the shadow will be invisible unless the default <code>shadow.col</code> is changed).
<code>nn.adj</code>	Horizontal position of the node label text. Default <code>.5</code> .
<code>nn.yshift</code>	Vertical position of the node numbers, in character heights relative to their default positions. Default <code>0</code> .
<code>nn.space</code>	Horizontal space to the box border on each side of the node number text, in character widths. Default <code>.8</code> .
<code>nn.yspace</code>	Vertical space to the box border above and below the node number text, in character heights. Default <code>.5</code> .
<code>under.ygap</code>	Applies if text is plotted under the box (i.e. if <code>under=TRUE</code> or there is a double newline in <code>prefix</code> or <code>suffix</code>). Vertical gap (in char heights) between the lower edge of the box and the top of the text under the box.
<code>yesno.yshift</code>	Vertical position of "yes" and "no" in character heights relative to their default position. Default <code>0</code> . Applies only when <code>yesno=TRUE</code> .

<code>ygap</code>	Minimum vertical gap between boxes, in character heights. Default <code>gap/2</code> .
<code>xcompact</code>	If TRUE (the default) and there is too much white space, automatically change <code>xlim</code> to compact the entire tree horizontally. This usually only activates for small trees. (The <code>xcompact</code> and <code>ycompact</code> arguments compact the tree as a whole, whereas the <code>compress</code> and <code>ycompress</code> arguments move parts of the tree into available space.)
<code>ycompact</code>	If TRUE (the default) and there is too much vertical space, automatically change <code>ylim</code> to compact the entire tree vertically.
<code>xcompact.ratio</code>	Default <code>.8</code> . Applies only when <code>xcompact=TRUE</code> . The maximum possible without overplotting is 1, but compacting by <code>.8</code> usually gives more pleasing spacing (it gives more space).
<code>min.inter.height</code>	Default 4. Applies only when <code>ycompact=TRUE</code> . Minimum height (in units of character height) between the lowest label in a layer and the highest label in the layer below it.
<code>max.auto.cex</code>	Clamp the maximum automatically calculated <code>cex</code> at this value. Default 1, meaning never expand <code>cex</code> , only contract when necessary.
<code>min.auto.cex</code>	Default <code>.15</code> . Never downscale to less than this when automatically calculating <code>cex</code> , even if overplotted labels result. (The graph layout algorithm is unstable with <code>cex</code> 's below <code>0.15</code> – meaning that the automatic type size may be smaller than necessary.)
<code>ycompress.cex</code>	Default <code>.7</code> . Applies only when <code>ycompress=TRUE</code> . Apply the <code>ycompress</code> algorithm if the initial automatically calculated <code>cex</code> is less than this. The idea is that we don't want to shift if we get an acceptable <code>cex</code> without shifting. Make <code>Inf</code> to always attempt shifting.
<code>accept.cex</code>	Accept shifting only if it causes at least this much improvement in <code>cex</code> (because we don't want to shift if it gives only a small improvement in <code>cex</code>). Default <code>1.1</code> i.e. require at least a 10% improvement. Use <code>0</code> to always accept shifts and <code>Inf</code> to never accept (or use <code>ycompress=FALSE</code>).
<code>shift.amounts</code>	Default <code>c(1.5, 2, 3)</code> . For <code>ycompress</code> , choose the best <code>cex</code> yielded by shifting nodes by these amounts, in multiples of the box heights (after initial scaling).
<code>Fallen.yspace</code>	Extra space for fallen leaves. Default <code>.1</code> , meaning allow 10% of the vertical space for the fallen leaves. (The name <code>Fallen.yspace</code> uses upper case to avoid partial matching with <code>fallen.leaves</code> .)
<code>boxes.include.gap</code>	Default FALSE. Include <code>gap</code> and <code>ygap</code> when drawing the boxes, for debugging purposes. (To draw the boxes, see <code>box.col</code> , <code>border.col</code> , <code>split.box.col</code> , and <code>split.border.col</code> .) This argument only affects the way the boxes are drawn, not the graph layout algorithm in any way. With the optimum <code>cex</code> at least one pair of boxes displayed in this manner will just touch (but none will overlap).
<code>legend.x</code>	Applies only to models with a multilevel class response (not binary or anova models, for which no legend is drawn). Horizontal position of the legend. Typically a value between <code>0</code> and <code>1</code> , although values beyond those limits are often useful. Default is NULL meaning automatically position the legend (assuming there is enough space). Use NA for no legend.

legend.y	Like legend.x but for the vertical position of the legend.
legend.cex	Like legend.x but for the relative size of the legend text. Default is 1.
...	Extra par arguments. Only the “important” par arguments are supported. Note that arguments like col apply only to the <i>node</i> labels. To affect the split labels or branch lines, use split.col and branch.col instead.

Value

A list with the following components. With the default args most of these are calculated automatically.

obj	The rpart object. Identical to the x argument passed in unless snip was used.
snipped.nodes	The snipped nodes, NULL unless snip was used.
xlim, ylim	The graph limits.
x, y	The node coords.
branch.x, branch.y	The branch line coords.
labs	The node labels.
cex	The node label cex.
boxes	The coords of the boxes around the nodes.
split.labs	The split labels.
split.cex	The split label cex.
split.boxes	The coords of the boxes around the splits.

See Also

The package vignette [Plotting rpart trees with the rpart.plot package](#)

[rpart.plot](#)

Functions in the rpart package: [plot.rpart](#) [text.rpart](#) [rpart](#)

Examples

```
data(ptitanic)
tree <- rpart(survived ~ ., data=ptitanic, cp=.02)
# cp=.02 because want small tree for demo

old.par <- par(mfrow=c(2,2))
# put 4 figures on one page

prp(tree, main="default prp\n(type = 0, extra = 0)")

prp(tree, main="type = 4, extra = 6\nbox.palette = \"auto\"",
     type=4, extra=6, # label all nodes, show prob of second class
     box.palette="auto", # auto color the nodes based on the model type
     faclen=0) # faclen=0 to print full factor names
```

```

cols <- ifelse(tree$frame$yval == 1, "darkred", "green4")
           # green if survived

prp(tree, main="assorted arguments",
     extra=106,      # display prob of survival and percent of obs
     nn=TRUE,       # display the node numbers
     fallen.leaves=TRUE, # put the leaves on the bottom of the page
     shadow.col="gray", # shadows under the leaves
     branch.lty=3,   # draw branches using dotted lines
     branch=.5,     # change angle of branch lines
     faclen=0,      # faclen=0 to print full factor names
     trace=1,       # print the automatically calculated cex
     split.cex=1.2, # make the split text larger than the node text
     split.prefix="is ", # put "is " before split text
     split.suffix="?", # put "?" after split text
     col=cols, border.col=cols, # green if survived
     split.box.col="lightgray", # lightgray split boxes (default is white)
     split.border.col="darkgray", # darkgray border on split boxes
     split.round=.5) # round the split box corners a tad

# compare to the plotting functions in the rpart package
plot(tree, uniform=TRUE, compress=TRUE, branch=.2)
text(tree, use.n=TRUE, cex=.8, xpd=NA) # cex is a guess, depends on your window size
title("compare to the plotting functions\nin the rpart package", cex.sub=.8)

par(old.par)

```

ptitanic

Titanic data with passenger names and other details removed.

Description

Titanic data with passenger names and other details removed.

Format

A data frame with 1046 observations on 6 variables.

pclass	passenger class, unordered factor: 1st 2nd 3rd
survived	factor: died or survived
sex	unordered factor: male female
age	age in years, min 0.167 max 80.0
sibsp	number of siblings or spouses aboard, integer: 0...8
parch	number of parents or children aboard, integer: 0...6

Source

The dataset was compiled by Frank Harrell and Robert Dawson:

<http://biostat.mc.vanderbilt.edu/wiki/pub/Main/DataSets/titanic.html>.

See also:

<http://biostat.mc.vanderbilt.edu/twiki/pub/Main/DataSets/titanic3info.txt>.

For this version of the Titanic data, passenger details were deleted, survived was cast as a factor, and the name changed to ptitanic to minimize confusion with other versions.

In this data the crew are conspicuous by their absence.

Contents of ptitanic:

	pclass	survived	sex	age	sibsp	parch
1	1st	survived	female	29.000	0	0
2	1st	survived	male	0.917	1	2
3	1st	died	female	2.000	1	2
4	1st	died	male	30.000	1	2
5	1st	died	female	25.000	1	2
...						
1309	3rd	died	male	29.000	0	0

How ptitanic was built:

```
load("titanic3.sav") # from Dr. Harrell's web site
# discard name, ticket, fare, cabin, embarked, body, home.dest
ptitanic <- titanic3[,c(1,2,4,5,6,7)]
# change survived from integer to factor
ptitanic$survived <- factor(ptitanic$survived, labels=c("died", "survived"))
save(ptitanic, file="ptitanic.rda")
```

This version of the data differs from `etitanic` in the `earth` package in that here `survived` is a factor (not an integer) and `age` has some NAs.

Examples

```
data(ptitanic)
summary(ptitanic)
# main indicator of missing data is 3rd class esp. with many children
obs.with.nas <- rowSums(is.na(ptitanic)) > 0
prp(rpart(obs.with.nas~., data=ptitanic, method="class"),
    main="observations with missing data", extra=7)
```

<code>rpart.plot</code>	<i>Plot an rpart model. A simplified interface to the prp function.</i>
-------------------------	---

Description

Plot an `rpart` model, automatically tailoring the plot for the model's response type.

For an overview, please see the package vignette [Plotting rpart trees with the rpart.plot package](#).

This function is a simplified front-end to `prp`, with only the most useful arguments of that function, and with different defaults for some of the arguments.

The different defaults for the `extra` and `col.palette` arguments mean that this function automatically creates a colored plot suitable for the type of model (whereas `prp` by default creates a minimal plot). In detail the different defaults are:

	<code>rpart.plot</code>	<code>prp</code>	
type	2	0	
extra	"auto"	0	
fallen.leaves	TRUE	FALSE	
varlen	0	-8	
faclen	0	3	
box.palette	"auto"	0	

The function `rpart.plot.version1` is compatible with old versions of this function and has the same defaults as `prp`.

Usage

```
rpart.plot(x=stop("no 'x' arg"),
  type=2, extra="auto",
  under=FALSE, fallen.leaves=TRUE,
  digits=2, varlen=0, faclen=0,
  cex=NULL, tweak=1,
  snip=FALSE,
  box.palette="auto", shadow.col=0,
  ...)
```

Arguments

To start off, look at the arguments `x`, `type` and `extra`. Just those arguments will suffice for many users. If you don't want a colored plot, use `box.palette=0`.

An `rpart` object. The only required argument.

`type`

Type of plot. Possible values:

0 Draw a split label at each split and a node label at each leaf.

1 Label all nodes, not just leaves. Similar to `text.rpart`'s `all=TRUE`.

2 Default. Like 1 but draw the split labels below the node labels. Similar to the plots in the CART book.

3 Draw separate split labels for the left and right directions.

4 Like 3 but label all nodes, not just leaves. Similar to `text.rpart`'s `fancy=TRUE`. See also `clip.right.labs`.

extra

Display extra information at the nodes. Possible values:

"auto" (case insensitive) Default.

Automatically select a value based on the model type, as follows:

extra=106 class model with a binary response

extra=104 class model with a response having more than two levels

extra=100 other models

0 No extra information.

1 Display the number of observations that fall in the node (per class for class objects; prefixed by the number of events for poisson and exp models). Similar to `text.rpart`'s `use.n=TRUE`.

2 Class models: display the classification rate at the node, expressed as the number of correct classifications and the number of observations in the node. Poisson and exp models: display the number of events.

3 Class models: misclassification rate at the node, expressed as the number of incorrect classifications and the number of observations in the node.

4 Class models: probability per class of observations in the node (conditioned on the node, sum across a node is 1).

5 Class models: like 4 but do not display the fitted class.

6 Class models: the probability of the second class only. Useful for binary responses.

7 Class models: like 6 but do not display the fitted class.

8 Class models: the probability of the fitted class.

9 Class models: the probabilities times the fraction of observations in the node (the probability relative to all observations, sum across all leaves is 1).

+100 Add 100 to any of the above to also display the percentage of observations in the node. For example `extra=101` displays the number and percentage of observations in the node. Actually, it's a weighted percentage using the weights passed to `rpart`.

Note 1: Unlike `text.rpart`, by default `prp` uses its own routine for generating node labels (not the function attached to the object). See the `node.fun` argument of `prp`.

Note 2: The `extra` argument has special meaning for `mvpart` objects. See the Appendix to this package's vignette.

under

Applies only if `extra > 0`. Default FALSE, meaning put the extra text *in* the box. Use TRUE to put the text *under* the box.

fallen.leaves

Default TRUE to position the leaf nodes at the bottom of the graph. It can be helpful to use FALSE if the graph is too crowded and the text size is too small.

digits	<p>The number of significant digits in displayed numbers. Default 2. If 0, use <code>getOption("digits")</code>.</p> <p>Details: Numbers from 0.001 to 9999 are printed without an exponent (and the number of digits is actually only a suggestion, see <code>format</code> for details). Numbers out that range are printed with an “engineering” exponent (a multiple of 3).</p>
varlen	<p>Length of variable names in text at the splits (and, for class responses, the class in the node label). Default 0, meaning display the full variable names. Possible values:</p> <p>0 use full names (default).</p> <p>greater than 0 call <code>abbreviate</code> with the given <code>varlen</code>.</p> <p>less than 0 truncate variable names to the shortest length where they are still unique, but never truncate to shorter than <code>abs(varlen)</code>.</p>
facLen	<p>Length of factor level names in splits. Default 0, meaning display the full factor names. Possible values are as <code>varlen</code> above, except that for back-compatibility with <code>text.rpart</code> the special value 1 means represent the factor levels with alphabetic characters (a for the first level, b for the second, etc.).</p>
cex	<p>Default NULL, meaning calculate the text size automatically.</p>
tweak	<p>Adjust the (possibly automatically calculated) <code>cex</code>. Default 1, meaning no adjustment. Use say <code>tweak=1.2</code> to make the text 20% larger. However, since font sizes are discrete the <code>cex</code> you ask for may not be exactly the <code>cex</code> you get. And a small change to <code>tweak</code> may not actually change the type size, or change it more than you want.</p>
snip	<p>Default FALSE. Set TRUE to interactively trim the tree with the mouse. See the <code>package vignette</code> (or just try it).</p>
box.palette	<p>Palette for coloring the node boxes based on the fitted value. This is a vector of <code>colors</code>, for example <code>box.palette=c("green", "green2", "green4")</code>. Small fitted values are displayed with colors at the start of the vector; large values with colors at the end.</p> <p>The special value <code>box.palette=0</code> (default for <code>prp</code>) uses the background color (typically white).</p> <p>The special value <code>box.palette="auto"</code> (default for <code>rpart.plot</code>, case insensitive) automatically selects a predefined palette based on the type of model.</p> <p>Otherwise specify a predefined palette e.g. <code>box.palette="Grays"</code> for the predefined gray palette (a range of grays). The predefined palettes are (see the <code>show.prp.palettes</code> function):</p> <p>Grays Greys Greens Blues Browns Oranges Reds Purples Gy Gn Bu Bn Or Rd Pu (alternative names for the above palettes) BuGn GnRd BuOr etc. (two-color diverging palettes: any combination of two of the above palettes) RdYlGn GnYlRd (three color palettes)</p> <p>Prefix the palette name with “-” to reverse the order of the colors e.g. <code>box.palette="-auto"</code> or <code>box.palette="-Grays"</code>.</p>

shadow.col Color of the shadow under the boxes. Default \emptyset , no shadow. Try "gray" or "darkgray".

... Extra arguments passed to [prp](#) and the plotting routines. Any of [prp](#)'s arguments can be used.

Value

The returned value is identical to that of [prp](#).

Author(s)

Stephen Milborrow, borrowing heavily from the [rpart](#) package by Terry M. Therneau and Beth Atkinson, and the R port of that package by Brian Ripley.

See Also

The package vignette [Plotting rpart trees with the rpart.plot package](#)

[prp](#)

[rpart.plot.version1](#)

Functions in the rpart package: [plot.rpart](#) [text.rpart](#) [rpart](#)

Examples

```
old.par <- par(mfrow=c(2,2))                      # put 4 figures on one page

data(ptitanic)

#-----

binary.model <- rpart(survived ~ ., data=ptitanic, cp=.02)
                                                  # cp=.02 for small demo tree

rpart.plot(binary.model,
          main="titanic survived\n(binary response)")

rpart.plot(binary.model,
          box.palette="Grays",                  # override default GnBu palette
          main="titanic survived\nbox.palette = \"Grays\"")

#-----

anova.model <- rpart(Mileage ~ ., data=cu.summary)

rpart.plot(anova.model,
          shadow.col="gray",                    # add shadows just for kicks
          main="miles per gallon\n(continuous response)\n")

#-----
```

```
multi.class.model <- rpart(Reliability ~ ., data=cu.summary)

rpart.plot(multi.class.model,
           main="vehicle reliability\n(multi class response)")

par(old.par)
```

rpart.plot.version1 *Plot an rpart model (old version).*

Description

This function is provided for backward compatibility with old versions of `rpart.plot` (prior to version 2.0.0, June 2016).

By default this function creates a minimal unadorned plot, whereas `rpart.plot` creates a colored plot automatically tailored to the model's type.

Actually, the only difference between the two functions is that some arguments have different defaults. This function has the same defaults as `prp`. See the `rpart.plot` help page for a table showing the different defaults.

Usage

```
rpart.plot.version1(x=stop("no 'x' arg"),
                   type=0, extra=0,
                   under=FALSE, fallen.leaves=FALSE,
                   digits=2, varlen=-8, faclen=3,
                   cex=NULL, tweak=1,
                   snip=FALSE,
                   box.palette=0, shadow.col=0,
                   ...)
```

Arguments

<code>x</code>	Same as <code>rpart.plot</code> .
<code>type</code>	Default 0. Same as <code>rpart.plot</code> .
<code>extra</code>	Default 0, otherwise the same as <code>rpart.plot</code> (which has a default of "auto").
<code>under</code>	Default FALSE. Same as <code>rpart.plot</code> .
<code>fallen.leaves</code>	Default FALSE, otherwise the same as <code>rpart.plot</code> (which has a default of TRUE).
<code>digits</code>	Default 2. Same as <code>rpart.plot</code> .
<code>varlen</code>	Default -8, otherwise the same as <code>rpart.plot</code> (which has a default of 0).
<code>faclen</code>	Default 3, otherwise the same as <code>rpart.plot</code> (which has a default of 0).
<code>cex</code>	Default NULL. Same as <code>rpart.plot</code> .
<code>tweak</code>	Default 1. Same as <code>rpart.plot</code> .

snip	Default FALSE. Same as <code>rpart.plot</code> .
box.palette	Default 0, otherwise the same as <code>rpart.plot</code> (which has a default of "auto").
shadow.col	Default 0. Same as <code>rpart.plot</code> .
...	Same as <code>rpart.plot</code> .

Value

The returned value is identical to that of `prp`.

Examples

```
data(ptitanic)
tree <- rpart(survived ~ ., data=ptitanic, cp=.02)
      # cp=.02 because want small tree for demo

old.par <- par(mfrow=c(2,2))
      # put 4 figures on one page

rpart.plot.version1(tree, main="default rpart.plot.version1\n(type = 0, extra = 0)")

rpart.plot.version1(tree, main="type = 4, extra = 6",
                    type=4, extra=6, faclen=0) # faclen=0 to print full factor names

rpart.plot.version1(tree, main="extra = 106, under = TRUE",
                    extra=106, under=TRUE, faclen=0)

# compare to the plotting functions in the rpart package
plot(tree, uniform=TRUE, compress=TRUE, branch=.2)
text(tree, use.n=TRUE, cex=.8, xpd=NA) # cex is a guess, depends on your window size
title("compare to the plotting functions\nin the rpart package", cex.sub=.8)

par(old.par)
```

show.prp.palettes *Show the built-in prp palettes.*

Description

Display a diagram showing the built-in palettes accepted by `prp`'s `box.palette` argument.

Usage

```
show.prp.palettes()
```

Examples

```
show.prp.palettes()
```

Index

*Topic **CART**

prp, [2](#)
rpart.plot, [16](#)
rpart.plot.version1, [20](#)

*Topic **datasets**

ptitanic, [14](#)

*Topic **partitioning**

prp, [2](#)
rpart.plot, [16](#)
rpart.plot.version1, [20](#)

*Topic **recursive**

prp, [2](#)
rpart.plot, [16](#)
rpart.plot.version1, [20](#)

*Topic **rpart**

prp, [2](#)
rpart.plot, [16](#)
rpart.plot.version1, [20](#)

*Topic **tree**

prp, [2](#)
rpart.plot, [16](#)
rpart.plot.version1, [20](#)

abbreviate, [5](#), [18](#)

alpha, [7](#)

colors, [6](#), [18](#)

earth, [15](#)

etitanic, [15](#)

format, [5](#), [18](#)

frame, [4](#)

getOption, [5](#), [18](#)

par, [13](#)

plot.rpart, [2](#), [13](#), [19](#)

prp, [2](#), [2](#), [16](#), [19–21](#)

ptitanic, [14](#)

rpart, [2](#), [3](#), [13](#), [16](#), [19](#)

rpart.plot, [2](#), [13](#), [16](#), [16](#), [20](#), [21](#)

rpart.plot.version1, [16](#), [19](#), [20](#)

show.prp.palettes, [6](#), [18](#), [21](#)

text, [10](#)

text.rpart, [2](#), [5](#), [13](#), [18](#), [19](#)