Package 'Snowball'

September 19, 2011

R topics documented:

SnowballStemmer *R/Weka Snowball Stemmer*

Description

R interface to Weka's Snowball stemmers.

Usage

```
SnowballStemmer(x, control = NULL)
```

Arguments

x   a character vector with words to be stemmed.

control   an object of class [Weka_control](), or a character vector of control options, or
                NULL (default). Available options can be obtained on-line using the Weka Option
                Wizard [WOW](), or the Weka documentation.

Details

The Snowball stemmers contain the Porter stemmer and several other stemmers for different lan-
guages. See <http://snowball.tartarus.org/> for more information.

SnowballStemmer is an interface to Weka's wrapper classes for the Java version of the Snowball
stemmers. The corresponding jar cannot be included in package **RWeka** due to license restrictions,
and hence is made available via the separate package **Snowball**.

The Omegahat package **Rstem** provides an R interface to a C version of Porter's word stemming
algorithm.

Value

A character vector with the stemmed words.

See Also

Other R interfaces to Weka stemmers ([RWeka_stemmers]())

Examples

```
## Test the supplied vocabulary for the default stemmer ('porter'):
source <- readLines(system.file("words", "porter","voc.txt",
                                package = "Snowball"))
result <- SnowballStemmer(source)
target <- readLines(system.file("words", "porter", "output.txt",
                                package = "Snowball"))
## Any differences?
any(result != target)
```

Index