cleangeo - Cleaning Geometries from Spatial Objects in R

Emmanuel Blondel
emmanuel.blondel1@gmail.com

CEO - International Consultant

September 30, 2015

Outline

1 Introduction

2 Using cleangeo

3 Conclusions & Perspectives

cleangeo
R spatial data processing

R provides a lot of facilities to read and manipulate spatial data. We can mention some packages such as:

- `sp`, which provides a set of classes and methods to handle spatial objects
- `maptools`, a set of tools for reading and handling Spatial Objects
- `rgdal`, which provides an interface to the well-known Geospatial Data Abstraction Library (GDAL)
- `rgeos`, which provides an interface to the widely used Geometry Engine, Open Source (GEOS) library.

The main problem comes when spatial objects did not have valid geometries, or expose different types of geometry errors (e.g. orphaned holes), which prevents any easy spatial data processing.

cleangeo

A package for cleaning geometries from spatial objects in R

`cleangeo` was built in order:

- to facilitate handling and catching rgeos geometry issues
- to provide an utility to clean the spatial objects.

`cleangeo` appears to a useful tool to check, and eventual clean the data prior any spatial data processing.

Outline

1 Introduction

2 Using cleangeo

3 Conclusions & Perspectives

cleangeo
Usage - Installing cleangeo

`cleangeo` can be installed:

- from CRAN

```
R> install.packages("cleangeo")
```

- from Github, for latest updates (requires `devtools` package)

```
R> require(devtools)
R> install_github("eblondel/cleangeo")
```

Load `cleangeo` in R using:

```
R> require(cleangeo)
```

cleangeo

Usage - Identify geometry issues in spatial objects

Let's load some sample data in R:

```
R> require(maptools)
R> file <- system.file("extdata", "example.shp", package = "cleangeo")
R> sp <- readShapePoly(file)
```

cleangeo first allows to inspect spatial objects, and detect eventual issues:

```
R> sp.report <- clgeo_CollectionReport(sp)
R> sp.summary <- clgeo_SummaryReport(sp.report)
```

```
             type     valid                issue_type
 rgeos_validity:2   Mode :logical   GEOM_VALIDITY:2
 NA's           :1   FALSE:2         NA's          :1
                     TRUE :1
                     NA's :0
```

The output indicates that 2 spatial objects have a problem of geometry validity.

cleangeo

Usage - Clean geometries spatial objects

`cleangeo` allows you to clean spatial objects:

```
R> sp.fixed <- clgeo_Clean(sp)
```

Let's analyse the new spatial objects:

```
R> sp.fixed.report <- clgeo_CollectionReport(sp.fixed)
R> sp.fixed.summary <- clgeo_SummaryReport(sp.fixed.report)
```

```
   type     valid          issue_type
 NA's:3   Mode:logical   NA's:3
          TRUE:3
          NA's:0
```

And that's it! Now your spatial objects have clean geometries, you can safely do some spatial data processing on it.

Outline

1 Introduction

2 Using cleangeo

3 Conclusions & Perspectives

cleangeo
Conclusions and Perspectives

Although `cleangeo` is at its early stages, it has proved to be a real benefit in the field of spatial data processing, providing a **quick** and **simple** way to check geometry issues and clean spatial objects.

`cleangeo` is currently focused on cleaning `SpatialPolygons` objects. A direct perspective is to extend it to other `Spatial*` objects.

Any suggestion from users is welcome to strenghten the `cleangeo` package.

cleangeo
Looking for Sponsors

`cleangeo` was initially born from some assistance provided to users that were facing issues in processing spatial data in R (see the original post). The prototype package was made available on Github on a voluntary basis in support to this users request.

To guarantee the sustainability of `cleangeo`, we are seeking for **fundings**, throught *sponsoring* or *donations* to:

- **implement, test, validate and release improvements**
- **guarantee a quality maintenance** of the package
- **provide support** to users and institutions that may take advantage of `cleangeo`

If you are interesting in supporting `cleangeo`, please do not hesitate to contact me!

cleangeo
on the web

- on Github:
  - source code: https://github.com/eblondel/cleangeo
  - online documentation: https://github.com/eblondel/cleangeo/blob/master/vignettes/quickst
- on the Comprehensive R Archive Network (CRAN): http://cran.r-project.org/package=cleangeo