Data Transformation Exercise in R: Converting a nested list into dataframe

Mert Nuhoglu

July 21, 2015

Problem Definition

- Conversion of a list into a data frame (or data table)
- Lists: Hierarchical
- Data Frames: Flat

Source Data Structure

- To reproduce the list use the following code

```
f = list(list(NA, "false", "5"), list(NA, "false", "3"))
```

Structure of List

```
str(f)
```

```
## List of 2
##  $ :List of 3
##   ..$ : logi NA
##   ..$ : chr "false"
##   ..$ : chr "5"
##  $ :List of 3
##   ..$ : logi NA
##   ..$ : chr "false"
##   ..$ : chr "3"
```

Target Data Structure

- A dataframe of two rows and three columns:

	X1	X2	X3
1	false	5	2
false	3		

Alternative 1: Can we use rbind or rbindlist?

```
do.call(rbind, f)
```

```
##      [,1] [,2]    [,3]
## [1,] NA   "false" "5"
## [2,] NA   "false" "3"
```

```
rbindlist(f)
```

```
##    V1    V2 V3
## 1: NA false  5
## 2: NA false  3
```

Converting matrix to data.table

```
do.call(rbind, f) %>%
    as.data.table
```

```
##    V1    V2 V3
## 1: NA false  5
## 2: NA false  3
```

```
do.call(rbind, f) %>%
    as.data.frame
```

```
##   V1    V2 V3
## 1 NA false  5
## 2 NA false  3
```

List of dataframe

```r
g = list(
    data.frame(a=NA, b="false", c="5"),
    data.frame(a=NA, b="false", c="3")
)
str(g)
```

```
## List of 2
##  $ :'data.frame':    1 obs. of  3 variables:
##   ..$ a: logi NA
##   ..$ b: Factor w/ 1 level "false": 1
##   ..$ c: Factor w/ 1 level "5": 1
##  $ :'data.frame':    1 obs. of  3 variables:
##   ..$ a: logi NA
##   ..$ b: Factor w/ 1 level "false": 1
##   ..$ c: Factor w/ 1 level "3": 1
```

Can we use rbind or rbindlist? 2

```
do.call(rbind, g)
```

```
##    a     b c
## 1 NA false 5
## 2 NA false 3
```

```
rbindlist(g)
```

```
##     a     b c
## 1: NA false 5
## 2: NA false 3
```

Does rbind/rbindlist work always?

```
f2 = list(
    list(list(7, 2), "false", "5"),
    list(list(1, 3), "false", "3")
)
```

```
do.call(rbind, f2) %>%
    as.data.frame
```

```
##      V1    V2 V3
## 1 7, 2 false  5
## 2 1, 3 false  3
```

```
rbindlist(f2)
```

```
Error in rbindlist
```

rbind with deeply nested list

```
d2 = do.call(rbind, f2) %>%
    as.data.frame
str(d2)
```

```
## 'data.frame':    2 obs. of  3 variables:
##  $ V1:List of 2
##   ..$ :List of 2
##   .. ..$ : num 7
##   .. ..$ : num 2
##   ..$ :List of 2
##   .. ..$ : num 1
##   .. ..$ : num 3
##  $ V2:List of 2
##   ..$ : chr "false"
##   ..$ : chr "false"
##  $ V3:List of 2
##   ..$ : chr "5"
##   ..$ : chr "3"
```