

ToreeScala

September 10, 2018

0.1 Using external framework through Magics

- Add the jars with their compile dependencies; relaunch the kernel, then the cell

```
In [1]: %Addjar -f https://dl.bintray.com/spark-packages/maven/graphframes/graphframes/0.6.0-spark2.2-s_2.11.jar
        %Addjar -f http://central.maven.org/maven2/com/typesafe/scala-logging/scala-logging-api_2.11-2.1.2.jar
        %Addjar -f http://central.maven.org/maven2/com/typesafe/scala-logging/scala-logging-slf4j_2.11-2.1.2.jar
```

```
Starting download from https://dl.bintray.com/spark-packages/maven/graphframes/graphframes/0.6.0-spark2.2-s_2.11.jar
Finished download of graphframes-0.6.0-spark2.2-s_2.11.jar
Starting download from http://central.maven.org/maven2/com/typesafe/scala-logging/scala-logging-api_2.11-2.1.2.jar
Finished download of scala-logging-api_2.11-2.1.2.jar
Starting download from http://central.maven.org/maven2/com/typesafe/scala-logging/scala-logging-slf4j_2.11-2.1.2.jar
Finished download of scala-logging-slf4j_2.11-2.1.2.jar
```

```
In [9]: import org.graphframes.examples.Graphs
        val g = Graphs.friends
```

```
g = GraphFrame(v:[id: string, name: string ... 1 more field], e:[src: string, dst: string ... 1 more field])
```

```
Out[9]: GraphFrame(v:[id: string, name: string ... 1 more field], e:[src: string, dst: string ... 1 more field])
```

```
In [10]: // Query: Get in-degree of each vertex.
         g.inDegrees.show()
```

```
[Stage 7:=====> (87 + 2) / 100] +---+-----+
| id|inDegree|
+---+-----+
|  f|        1|
|  e|        1|
|  d|        1|
|  c|        2|
|  b|        2|
|  a|        1|
+---+-----+
```

```
In [11]: // Query: Count the number of "follow" connections in the graph.
         g.edges.filter("relationship = 'follow'").count()
```

```
Out[11]: 4
```

```
In [12]: // Run PageRank algorithm, and show results.
         val results = g.pageRank.resetProbability(0.01).maxIter(3).run()
         results.vertices.select("id", "pagerank").show()
```

```
[Stage 85:=====>      (184 + 2) / 200] +---+-----+
| id|                pagerank|
+---+-----+
| g| 0.011647254575707158|
| c|  3.1654026206322796|
| f| 0.30279950083194684|
| b|  2.6002789933444266|
| a|  0.3114187603993345|
| e|  0.3056533693843595|
| d| 0.30279950083194684|
+---+-----+
```

```
results = GraphFrame(v:[id: string, name: string ... 2 more fields], e:[src: string, dst: string
```

```
Out[12]: GraphFrame(v:[id: string, name: string ... 2 more fields], e:[src: string, dst: string
```