CSC512/495 Course Project: Spark API Converter

This project was done as a course project for the CSC512 Compiler Construction course. Authors:

1. Brayden McDonald
2. Igibek Koishybayev

Get start

The package includes everything to build the project. Folder structure:

- **spark/**
  - App.java: main java class, which contains entry point to the program
  - **dataframe/**
    - **CustomDataframeListener.java**: java class which extends antlr generated *DataframeListener.java* class. Each function inside CustomDataframeListener is fired when appropriate node is visited during traversal.
    - **Dataframe.g4**: spark RDD API program grammar used by antlr to generate classes.
  - **dataset/**
    - **CustomDatasetListener.java**: java class which extends antlr generated *DatasetListener.java* class. Each function inside CustomDataframeListener is fired when appropriate node is visited during traversal.
    - **Dataset.g4**: spark RDD API program grammar used by antlr to generate classes.
- **tests/**: contains 6 test programs to test implemented convertor.
- **part0/**: contains 18 programs that are implemented as Part-0 programs

Building the program

1. You need to export CLASSPATH variable. You can do it by running `source classpath.sh`.
2. Build the program using `make` command
3. To run the program type: `java spark.App <absolute_path_to_file> <api>`.

Examples:

1. Run a program for a **tests/dataset_1.t** file `java spark.App tests/dataset_1.t dataset`
2. Run a program for a **tests/dataframe_1.t** file `java spark.App tests/dataframe_1.t dataframe`