

### Finger 3: Graph Coloring con Spark! (individual)

Tiempo Estimado: 1 a 4 horas.

Fecha Límite: Lunes 12 de Septiembre

Puntos: 5

El objetivo de este TP es aprovechar la capacidad de Spark de procesar en paralelo. Como no todos contamos con acceso a un cluster creamos este finger que puede ser simulado perfectamente en una instalación de Spark en una sola computadora.

El objetivo del TP es resolver el famoso problema de Graph-Coloring, en este problema el objetivo es asignar a cada nodo de un grafo un color de forma tal que no haya aristas entre nodos de igual color. El desafío es encontrar la cantidad mínima de colores con los cuales podemos encontrar una solución.

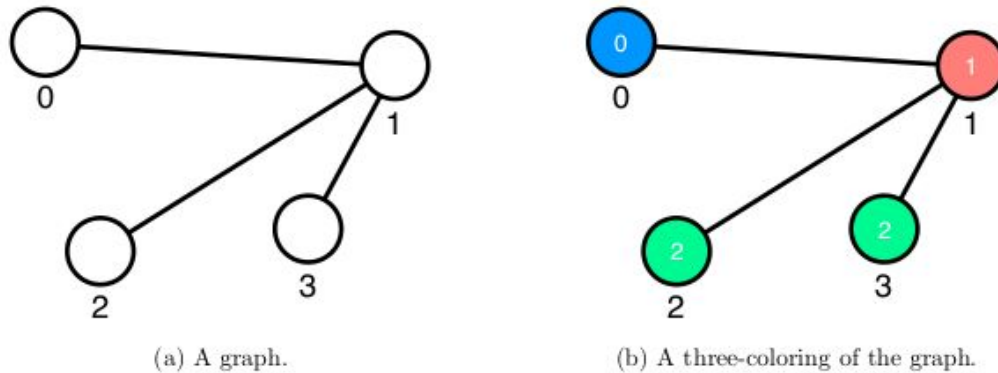


Figure 1: A Graph Coloring Example

Una forma de resolver este problema es por fuerza bruta, para una cierta cantidad de colores asignamos los colores al azar a los nodos y verificamos si cumple los requisitos. Si luego de "n" simulaciones no hemos logrado una solución incrementamos la cantidad de colores y repetimos el proceso. Están en libertad de probar algoritmos más interesantes si es que así lo desean.

El input para este finger es un conjunto de archivos con el siguiente formato:

```
N E
n_i n_j
n_i n_j
...
```

Ejemplo (4 nodos y 3 aristas) :

```
4 3
0 1
```

1 2  
1 3

La primera línea contiene dos números: N la cantidad de nodos y E la cantidad de aristas. Luego vienen E líneas en donde cada línea indica una arista entre dos nodos.

Una forma simple de distinguir la fila 1 de las demás en spark es:

```
input = sc.textFile(inputFile)

header = input.take(1)[0]

rows = input.filter(lambda line: line != header)
```

La salida debe ser un archivo que indique en la primera línea la cantidad de colores usados y luego una línea por cada nodo, es decir "n" líneas indicando en cada línea el color asignado a cada nodo. Por ejemplo:

3  
1  
1  
2  
3  
2

Esta solución representa una solución de 3 colores para un grafo de 5 nodos. El nodo "0" tiene color 1, el nodo "1" tiene color 1, el nodo "2" tiene color 2, el nodo 3 tiene color 3 y el nodo "4" tiene color 2.

Publicar en FB en el grupo de la materia:

- Cantidad de colores de la solución encontrada para cada archivo
- Link al código en Github para comprobar que la solución sea hecha con Apache Spark
- Comentarios que quieran hacer.

Puntos:

- 5,4,3,2 y varios de 1 punto a asignar de acuerdo a quienes tengan las mejores soluciones!