

## 1. Respuesta Breve : [jaccard\_mpi]

### 1.1. ¿Cuál o cuáles funciones de MPI consideró necesarias para la solución de este problema y explique cómo funciona(n) en su código?

Para solucionar este problema con MPI se implementaron 2 funciones.

```
MPI_Allgather(&max_jaccard_similarity, 1, MPI_DOUBLE, array_results, 1, MPI_DOUBLE, MPI_COMM_WORLD);
```

Primero, se utilizó un **Allgather**. El motivo de el Allgather es que va a existir en cada proceso un arreglo de resultados, por lo tanto ese arreglo tiene que tener los valores de max\_jaccard de cada proceso. Por ejemplo, si están ejecutando 2 procesos, el arreglo tendrá tamaño 2, y cada entrada equivale al max\_jaccard que obtuvo cada proceso. Esto es para lograr comparar los max\_jaccard de todos los procesos y saber, cuál es el verdadero max\_jaccard.

```
if (my_id != 0) {
    MPI_Recv(&message_received, 1 /* count*/, MPI_INT, my_id - 1 /*source*/, 123 /*message id*/, MPI_COMM_WORLD, &status);
}

if (ok) {
    cout << "\nMost similar text: \n\" << most_similar_text << "\"" << endl;
}

if (my_id == num_processes - 1)
    printf("\nJaccard similarity score: %.6f\n", max_jaccard_similarity);

if ( my_id < num_processes - 1 ) {
    MPI_Send(&message_sent, 1 /*count*/, MPI_INT, my_id + 1 /*dest*/, 123 /*message id*/, MPI_COMM_WORLD);
}

text_file.close();
delete[] array_results;
```

La otra función de MPI que se implementó fue una comunicación punto a punto entre todos los procesos, con el motivo para imprimir. Esto se vio necesario ya que en el caso que existan empates, se puedan imprimir las líneas empatadas de cada proceso en orden.

## 1.2. Explique cómo realizó la repartición de tareas entre los procesos.

```
max_jaccard_similarity = -1.0;
int i, j, k;
j = 0;
k = my_id;

while (getline(matrix_file, matrix_row) && getline(text_file, text_row)) {
    words_union = words_intersection = 0.0;

    if(j == k){
        for (i = 0; i < matrix_row.length(); ++i) {
            if (matrix_row[i] == '1' || test_vector[i] == '1') { // If the
                ++words_union;
            }
            if (matrix_row[i] == '1' && test_vector[i] == '1') { // If the
                ++words_intersection;
            }
        }

        jaccard_similarity = words_intersection / words_union; // Calculate

        if (jaccard_similarity > max_jaccard_similarity) { // Update highest
            max_jaccard_similarity = jaccard_similarity;
            most_similar_text = text_row;
        }
        k += num_processes;
    }
    j++;
}
```

Recordemos, que por ser MPI, los valores que cambian entre cada proceso son los id. Para hacer la repartición de tarea se hizo por un iterador. Como se puede apreciar en la imagen, j incrementa un valor en cada iteración del while, pero k está inicializado con el valor del id del proceso. Cuando se cumple la condición y se entra al if, a k se le suma la cantidad de procesos, para que de esta manera la repartición de tareas se haga de manera **cíclica**, y cada proceso tendrá un conjunto de trabajo diferente.

## 1.3. Casos de Prueba

```
You, seconds ago | 1 author (You)
1 quiero un monster burrito
2 quiero ir a taco bell
3 quiero llorar un rato
4 quiero llorar un rato
5 quiero llorar un rato
6 quiero llorar un rato
7 quiero ver a papa noel
8 quiero jugar bola
9 mi nombre es marco | You, se
```

```
mantofer@marco-hp-envy:~/Desktop/CI0117-2020-S2/Examenes/2/jaccard_mpi$ make  
mpic++ jaccard_mpi.cpp -o jaccard_mpi  
mantofer@marco-hp-envy:~/Desktop/CI0117-2020-S2/Examenes/2/jaccard_mpi$ mpiexec ./ja  
ext.txt" "../dataset_1/matrix.txt"
```

Evaluated text:

"quiero un monster burrito"

Most similar text:

"quiero llorar un rato"

Most similar text:

"quiero llorar un rato"

Most similar text:

"quiero llorar un rato"

Most similar text:

"quiero llorar un rato"

Jaccard similarity score: 0.333333

You, seconds ago | 1 author (You)

- 1 mi album favorito se llama to pimp a butterfly
- 2 el azul es mi color favorito
- 3 como se llama ese album
- 4 no entiendo que significa butterfly
- 5 quien es ud
- 6 como se llama ese color
- 7 yo si tengo hambre

You, seconds ago • Uncom

```
mantofer@marco-hp-envy:~/Desktop/CI0117-2020-S2/Examenes/2/jaccard_mpi$ mpiexec ./jaccard_mpi.py ext.txt" "../dataset_2/matrix.txt"
```

Evaluated text:

"mi album favorito se llama to pimp a butterfly"

Most similar text:

"como se llama ese album"

Jaccard similarity score: 0.272727

42 Iuste scelerisque a dui tempus atqueam. Duis hendrerit est nisl. Phasellus eget euismod odio, sed s  
43 Etiam in enim sed orci iaculis viverra sit amet vulputate odio. Cras ac pellentesque velit. Phasellus  
44 Phasellus sit amet odio scelerisque, feugiat diam sit amet, faucibus arcu. Class aptent taciti socios  
45 Nulla non ornare nibh. Nullam eu turpis eget massa eleifend eleifend. Integer in magna pulvinar, pulv  
46 Sed congue neque tellus, vel commodo elit bibendum a. Phasellus magna felis, pellentesque sed dui sed  
47 Nullam pellentesque a dui ut hendrerit. Aliquam iaculis tincidunt accumsan. Vestibulum a eleifend or  
48 Integer sollicitudin hendrerit ornare. Nunc sapien nisi, condimentum id tincidunt et, tempor a magna  
49 Quisque pharetra pulvinar mi vel mollis. Orci varius natoque penatibus et magnis dis parturient monte  
50 Donec fermentum congue quam, in semper nisi hendrerit ac. Aenean id erat a nunc ullamcorper sceleris  
51 Ut consequat, neque sit amet mattis hendrerit, est risus laoreet est, vitae pretium odio quam molest  
52 Nunc vel lacus eget orci luctus laoreet. Sed condimentum quam ante, at mollis lacus tempus id. Vestib  
53 Curabitur dignissim est risus, vel rutrum lorem auctor eget. Sed dignissim est sagittis, mollis purus  
54 Quisque at pretium mi. Duis condimentum turpis non porttitor bibendum. Aliquam auctor suscipit tempu  
55 Duis maximus velit ligula, varius dapibus est vehicula vel. Nam ut nunc id mi blandit aliquet bibend  
56 Integer vestibulum ligula vel posuere semper. Duis ultrices urna sit amet velit rutrum suscipit. Mori  
57 Nam nec tristique arcu, eget malesuada arcu. Sed consectetur, dolor vitae tristique dapibus, nunc to  
58 Sed luctus eleifend efficitur. Ut ac finibus est. In fringilla vel sem ut volutpat. Suspendisse luctu  
59 Aliquam ut leo malesuada, bibendum quam at, ullamcorper urna. Quisque imperdiet risus eget ipsum com  
60 Suspendisse porttitor lorem nec ullamcorper mollis. Integer ornare magna libero, auctor semper elit e  
61 Curabitur commodo dui in cursus condimentum. Sed vitae gravida orci. Nullam lobortis fermentum quam,  
62 Suspendisse dictum ac risus vel lobortis. Pellentesque nulla risus, blandit non nisl ac, pellentesque  
63 Aenean eu nulla ut eros blandit dictum. Suspendisse eget semper nisl. Aliquam vulputate tempus ligula  
64 Etiam mollis nulla vel volutpat efficitur. Sed at magna in est pellentesque consectetur nec sed leo.  
65 Nunc convallis consequat tellus, vitae auctor risus eleifend et. Nulla facilisi. Suspendisse dapibus  
66 Donec tincidunt maximus ante, efficitur viverra diam dapibus et. Integer purus mauris, tincidunt ac  
67 Quisque euismod dui at hendrerit ullamcorper. In dignissim, justo sit amet elementum gravida, sapien  
68 Suspendisse elementum laoreet tellus, quis consectetur purus interdum eget. Vestibulum imperdiet por  
69 Quisque iaculis nunc ac venenatis suscipit. Mauris eget fringilla tortor. Sed mi lectus, mollis quis  
70 Duis blandit euismod ipsum non rhoncus. Cras tempor, magna sed volutpat finibus, odio lacus suscipit

Evaluated text:

"Lorem ipsum dolor sit amet, consectetur adipiscing elit. In consequat rhoncus felis, et rhoncus tortor semper qui s. Vivamus quis mauris porta, aliquam nibh vitae, pharetra ipsum. Proin quam nulla, cursus vel velit quis, pulvina r pellentesque diam. Fusce non fermentum odio. Ut sodales porta ornare. Aliquam tincidunt congue arcu, a maximus n ulla consectetur a. Curabitur eget magna sit amet arcu facilisis viverra id fermentum dolor. Maecenas sit amet aug e ultrices, pellentesque eros nec, dictum erat. Cras pretium mi eget ante elementum ullamcorper at et orci."

Most similar text:  
 "In a sodales orci. Curabitur nec enim ornare nunc tincidunt suscipit. Lorem ipsum dolor sit amet, consectetur adi  
 piscin elit. Praesent consectetur rutrum mauris quis placerat. Maecenas vel neque scelerisque, auctor ligula et,  
 ornare elit. Nam et consequat odio. Quisque at diam bibendum, viverra ipsum vel, dictum erat. Integer fringilla sa  
 pien at dictum lobortis. Duis laoreet ultrices nisi eu pharetra. Suspendisse faucibus dictum erat sit amet consequ  
 at. Curabitur a mi ut ante placerat hendrerit sed id est. Etiam euismod turpis est. Proin vel auctor sapien. Morbi  
 eget quam aliquam, elementum magna sed, varius mi. Nam sodales quis risus sit amet rhoncus."

Jaccard similarity score: 0.247788