

# Práctica 4.b: Firefly Algorithm para el Problema del Agrupamiento con Restricciones:



# UNIVERSIDAD DE GRANADA

Luis González Romero, XXXXXXXXXX, [luisgonromero@correo.ugr.es](mailto:luisgonromero@correo.ugr.es)

Grupo 1: Miércoles 17:30-19:30

Escuela Técnica Superior de Ingeniería informática y Telecomunicaciones

27 de junio de 2020

# Práctica 4.b: Firefly Algorithm para el Problema del Agrupamiento con Restricciones

Memoria sobre la cuarta práctica de la asignatura Metaheurísticas  
cursada en la ETSIIT, UGR.

Luis González Romero, XXXXXXXXXX, luisgonromero@correo.ugr.es

Junio de 2020

# Índice

<b>1. Descripción del problema: Problema del Agrupamiento con Restricciones</b>	<b>5</b>
<b>2. Firefly Algorithm</b>	<b>6</b>
2.1. Descripción de la metaheurística . . . . .	6
2.2. Esquema de representación de soluciones . . . . .	7
2.3. Generación de la solución inicial . . . . .	7
2.4. Implementación FA . . . . .	8
<b>3. Hibridación FA-LS</b>	<b>10</b>
<b>4. Análisis de Resultados</b>	<b>11</b>
<b>5. Manual de uso</b>	<b>15</b>
<b>Referencias</b>	<b>15</b>

## Índice de tablas

1.	Resultados obtenidos por el algoritmo BL en el PAR con 10 % de restricciones	11
2.	Resultados obtenidos por el algoritmo BL en el PAR con 20 % de restricciones	11
3.	Resultados obtenidos por el algoritmo SA en el PAR con 10 % de restricciones	11
4.	Resultados obtenidos por el algoritmo SA en el PAR con 20 % de restricciones	11
5.	Resultados obtenidos por el algoritmo BMB en el PAR con 10 % de restricciones . . . . .	11
6.	Resultados obtenidos por el algoritmo BMB en el PAR con 20 % de restricciones . . . . .	12
7.	Resultados obtenidos por el algoritmo ILS en el PAR con 10 % de restricciones	12
8.	Resultados obtenidos por el algoritmo ILS en el PAR con 20 % de restricciones	12
9.	Resultados obtenidos por el algoritmo ILS-SA en el PAR con 10 % de restricciones . . . . .	12
10.	Resultados obtenidos por el algoritmo ILS-SA en el PAR con 20 % de restricciones . . . . .	12
11.	Resultados obtenidos por el algoritmo FA-LS en el PAR con 10 % de restricciones . . . . .	13
12.	Resultados obtenidos por el algoritmo FA-LS en el PAR con 20 % de restricciones . . . . .	13
13.	Resultados globales en el PAR con 10 % de restricciones . . . . .	13
14.	Resultados globales en el PAR con 20 % de restricciones . . . . .	13

## 1. Descripción del problema: Problema del Agrupamiento con Restricciones

El problema del *PAR* consiste en una generalización del agrupamiento clásico. Permite incorporar al proceso de agrupamiento un nuevo tipo de información: las restricciones. Dado un conjunto de datos  $X$  con  $n$  instancias, el problema consiste en encontrar una partición  $C = \{c_1, \dots, c_{1k}\}$  que minimice la desviación general y cumpla con las restricciones de instancia existentes en el conjunto  $R$ . Dada una pareja de instancias, se establece una restricción de tipo *Must-Link* (ML) si deben pertenecer al mismo grupo y de tipo *Cannot-Link* (CL) si no pueden pertenecer al mismo grupo.

Dada una partición  $C$ , se puede calcular el centroide  $\vec{\mu}_i$  asociado a cada grupo  $c_i$  como el vector promedio de las instancias de  $X$  que lo componen:

$$\vec{\mu}_i = \frac{1}{|c_i|} \sum_{\vec{x}_j \in c_i} \vec{x}_j$$

La distancia media intra-cluster  $\vec{c}_i$  será la media de las distancias de las instancias que lo conforman a su centroide:

$$\vec{c}_i = \frac{1}{|c_i|} \sum_{\vec{x}_j \in c_i} \|\vec{x}_j - \vec{\mu}_i\|_2$$

Definimos la desviación general de la partición  $C = \{c_1, \dots, c_{1k}\}$  como la media de las desviaciones intraccluster  $\vec{C}$ :

$$\vec{C} = \frac{1}{k} \sum_{c_i \in C} \vec{c}_i$$

Dada una partición  $C$  y los conjuntos de restricciones ML y CL, definimos infeasibility (“infactibilidad”) como el número de restricciones que  $C$  incumple.

Así, se puede formular como:

$$\text{Minimizar } f = \vec{C} + (\text{infeasibility} * \lambda)$$

siendo  $\lambda$  un parámetro de escalado para dar relevancia a la infeasibility. Así poder optimizar simultáneamente el número de restricciones incumplidas y la desviación general. Para asegurar que el factor infeasibility tiene la suficiente relevancia establecemos  $\lambda$  como el cociente entre la distancia máxima existente en el conjunto de datos y el número de restricciones presentes en el problema,  $|R|$ .

## 2. Firefly Algorithm

### 2.1. Descripción de la metaheurística

El FA algorithm es un algoritmo bioinspirado desarrollado por Xin-She Yang[1], en el que se reproduce el comportamiento de las luciérnagas. La luz que emiten las luciérnagas puede ser para atraer a posibles parejas (comunicación), posibles presas, entre otras. Sabemos que desde una determinada distancia  $r$  la intensidad lumínica desde su fuente se rige por la Ley de la inversa del cuadrado ( $I \propto 1/r^2$ ). Además, el aire absorbe parte de esta luz, decrementándose su intensidad cuanto mayor sea la distancia a este foco. Debido a estos dos factores, hacen que las luciérnagas sean visibles desde una cierta distancia, por la noche podrían ser cientos de metros que normalmente es suficiente para poder comunicarse entre ellas.

Para desarrollar el algoritmo correctamente podemos simplificarlo siguiendo las siguientes tres reglas:

- Todas las luciérnagas son unisex, por lo van a atraerse independientemente del sexo de estas.
- La atracción será proporcional a la luminosidad y ambas decrecen conforme la distancia aumente. Si no hay luciérnaga con mas intensidad lumínica, esta se moverá de forma aleatoria.
- La luminosidad vendrá dada por la función objetivo del problema, pudiendo tratarse de forma proporcional al valor de la función objetivo.

---

#### Firefly Algorithm

```
Objective function  $f(\mathbf{x})$ ,  $\mathbf{x} = (x_1, \dots, x_d)^T$   
Generate initial population of fireflies  $\mathbf{x}_i$  ( $i = 1, 2, \dots, n$ )  
Light intensity  $I_i$  at  $\mathbf{x}_i$  is determined by  $f(\mathbf{x}_i)$   
Define light absorption coefficient  $\gamma$   
while ( $t < \text{MaxGeneration}$ )  
  for  $i = 1 : n$  all  $n$  fireflies  
    for  $j = 1 : i$  all  $n$  fireflies  
      if ( $I_j > I_i$ ), Move firefly  $i$  towards  $j$  in  $d$ -dimension; end if  
      Attractiveness varies with distance  $r$  via  $\exp[-\gamma r]$   
      Evaluate new solutions and update light intensity  
    end for  $j$   
  end for  $i$   
  Rank the fireflies and find the current best  
end while  
Postprocess results and visualization
```

---

En el FA tenemos dos factores importantes: la variación de la intensidad lumínica y la atracción. Por simplicidad, podemos asumir que la atracción será determinada por la luminosidad, la cual está asociada a la función objetivo. Para problemas de optimización, la luminosidad en una localización concreta  $x$  podemos representarla como  $I \propto f(x)$ . No obstante, la luminosidad  $\beta$  de una luciérnaga es relativa, determinada por otras que la esten observando. La luz es absorbida por el medio y decrece con la distancia (como ya se ha comentado), pero por simplicidad como Yang comenta en su paper se usará  $f(x)$ .

Para la atracción, como es proporcional a la intensidad lumínica podemos calcularla de igual forma, determinada por la distancia  $r$  y el coeficiente de absorción  $\gamma$  para una determinada luciérnaga:  $\beta(r) = \beta_0 e^{-\gamma r^2}$  donde  $\beta_0$  es la atracción en  $r = 0$ . También puede calcularse de forma más rápida  $1/(1 + r^2)$ .

La distancia entre dos luciérnagas  $i$  y  $j$  en  $x_i$  y  $x_j$ , se corresponde a la distancia Cartesiana:

$$r_{ij} = ||x_i - x_j|| = \sqrt{\sum_{k=1}^d (x_{i,k} - x_{j,k})^2}$$

El movimiento de una luciérnaga  $i$  a otra mas atractiva (más luminosa)  $j$  viene dado por:

$$x_i = x_i + \beta(r) = \beta_0 e^{-\gamma r_{ij}^2} (x_j - x_i) + \alpha (rand - \frac{1}{2})$$

donde el segundo término se debe a la atracción y el tercero una componenta aleatoria, Yang recomienda definir  $\beta_0 = 1$ ,  $\gamma = 1$  y  $\alpha \in [0, 1]$ . Ahora el parámetro  $\gamma$  determina la velocidad de convergencia del algoritmo y el comportamiento de este.

## 2.2. Esquema de representación de soluciones

Para representar soluciones se sigue empleando un conjunto de valores de tamaño  $n_{instancias}$  para representarla solución, en el que cada instancia  $S_i$  queda asignada a un único cluster  $k$ .  $S = [S_0, S_1, ..., S_{n-1}]$

## 2.3. Generación de la solución inicial

Se genera un  $r_k$  (cluster aleatorio) y se añade a la estructura que guarda la solución. Cuando se obtiene la  $S$ , se comprueba que es válida (comprobando que al menos hay una ocurrencia de cada cluster) y si no lo es se repite el proceso hasta que se obtiene una factible.

---

**Algorithm 1** Generación de la solución inicial

---

```
procedure GET_SOLUCION_INICIAL()
  for  $i = 0$  to  $n\_instancias-1$  do
     $r\_k \leftarrow aleatorio(0, k - 1)$ 
     $c\_array[r\_k].append(i)$ 
     $S.append(r\_k)$ 
    while  $validar\_solucion\_actual() == \text{False}$  do
       $S.clear()$ 
       $c\_array.clear()$ 
      for  $i = 0$  to  $n\_instancias-1$  do
         $r\_k \leftarrow aleatorio(0, k - 1)$ 
         $c\_array[r\_k].append(i)$ 
         $S.append(r\_k)$ 
      end for
    end while
  end for
end procedure
```

---

## 2.4. Implementación FA

Para la implementación del algoritmo FA, cada luciérnaga la he representado como una solución(vector de asignaciones) y se mueven en base a estas asignaciones. En cuanto a la función de *attractiveness* he implementado y probado ambas, la gaussiana y la simplificada. Pero termine usando la simplificada ya que no diferían apenas las soluciones obtenidas y esta es menos costosa de calcular.

Para la generación inicial de las luciérnagas he reutilizado la forma de obtener el vector de asignaciones usado en el resto de prácticas. He usado la misma cantidad de luciérnagas que cromosomas en los genéticos, 50.

Miré algunos papers que usaban FA para clustering[2, 3] ya que no sabía si estaba enfocándolo bien. En estos papers se usan los centroides al tratarse el problema de clustering sin restricciones, pero yo decidí seguir con la representación que hemos usado en prácticas para poder manejar bien las restricciones. Al hacer el movimiento sobre el vector de asignaciones directamente, al realizar trazas veía cierta semejanza a los cruces de los genéticos, solo que no hereda directamente del padre sino que busca algo de calidad al acercarse a la luciérnaga de calidad encontrada en base a la distancia y su f.



---

**Algorithm 2** Algoritmo FA

---

```
procedure FA()
  inicializar_luciernagas_aleatorio()
  evaluar_luciernagas()
  evaluaciones  $\leftarrow$  0
  while evaluaciones < 100000 do
    for i = 0 to n_luciernagas-1 do
      mas_brillante  $\leftarrow$  TRUE
      for j = 0 to n_vecinos-1 do
        if intensidad_luciernaga[j] < intensidad_luciernaga[i] then
          moverse_hacia(i,j)
          mas_brillante  $\leftarrow$  FALSE
        end if
      end for
      if mas_brillante == FALSE then
        moverse_aleatoriamente()
      end if
    end for
  end while
end procedure
```

---

---

**Algorithm 3** Movimiento de una luciérnaga i hacia otra j

---

```
procedure MOVERSE_HACIA(LUCIERNAGA_I, LUCIERNAGA_J)
  copia  $\leftarrow$  luciernaga_i
  for n = 0 to n_instancias-1 do
    luciernaga_i[n]  $\leftarrow$  luciernaga_i[n] + attractiveness(luciernaga_i, luciernaga_j) *
    diferencia(luciernaga_i, luciernaga_j+aleatorio())
    luciernaga_i[n]  $\leftarrow$  luciernaga_i[n] % k
  end for
  if luciernaga_valida(luciernaga_i) == TRUE then
    actualizar_datos_luciernaga(luciernaga_i)
  else

    luciernaga_i  $\leftarrow$  copia
  end if
end procedure
```

---

### 3. Hibridación FA-LS

Para añadir explotación al algoritmo usé una soft LS como en la P2, pero no era suficiente para compensar la exploración tan grande del FA. Así que usé la LS original de la P1, y los resultados mejoraron aun más. Seguí un poco las trazas aplicando la LS a todas las luciérnagas, solo al 10 % y al 10 % de las mejores. La que más mejora consigue es al aplicar la LS al 10 % de los mejores, no ejecuté todas las semillas del resto de casos porque no mejoraban mucho y no tengo tiempo para dejar el ordenador ejecutando todo el rato que necesita. En los otros casos no mejora tanto como al aplicar la LS al 10 % de los mejores, quizás con un mejor tratamiento de los parámetros y buscar más equilibrio se podría mejorar aún más.

---

**Algorithm 4** Algoritmo FA

---

```
procedure FA()
  inicializar_luciernagas_aleatorio()
  evaluar_luciernagas()
  evaluaciones  $\leftarrow$  0
  ls  $\leftarrow$  0
  while evaluaciones < 100000 do
    ls  $\leftarrow$  ls + 1
    for i = 0 to n_luciernagas-1 do
      mas_brillante  $\leftarrow$  TRUE
      for j = 0 to n_vecinos-1 do
        if intensidad_luciernaga[j] < intensidad_luciernaga[i] then
          moverse_hacia(i,j)
          mas_brillante  $\leftarrow$  FALSE
        end if
      end for
      if mas_brillante == FALSE then
        moverse_aleatoriamente()
      end if
    end for
    if ls == frec_ls then
      ls  $\leftarrow$  0
      local_search()
    end if
  end while
end procedure
```

---

## 4. Análisis de Resultados

	Iris				Ecoli				Rand				Newthyroid			
	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T
Ejelución 1	0.57	0.00	0.57	2.07	23.08	70.00	24.96	172.58	0.76	0.00	0.76	1.42	10.80	114.00	14.97	9.14
Ejelución 2	0.57	0.00	0.57	2.23	23.05	94.00	25.56	213.69	0.76	0.00	0.76	2.17	10.86	97.00	14.41	8.95
Ejelución 3	0.57	0.00	0.57	1.79	23.19	68.00	25.02	204.21	0.76	0.00	0.76	1.31	13.83	6.00	14.05	9.41
Ejelución 4	0.57	0.00	0.57	2.15	21.25	122.00	24.52	183.51	0.76	0.00	0.76	1.69	10.81	122.00	15.27	7.60
Ejelución 5	0.57	0.00	0.57	1.62	21.24	121.00	24.49	127.18	0.76	0.00	0.76	1.50	10.88	98.00	14.47	9.45
Media	0.57	0.00	0.57	1.97	22.36	95.00	24.91	180.23	0.76	0.00	0.76	1.62	11.44	87.40	14.63	8.91

Tabla 1: Resultados obtenidos por el algoritmo BL en el PAR con 10 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T
Ejelución 1	0.57	0.00	0.57	3.09	21.89	152.00	23.93	167.08	0.76	0.00	0.76	2.33	14.29	0.00	14.29	10.85
Ejelución 2	0.57	0.00	0.57	2.97	18.69	402.00	24.08	148.31	0.76	0.00	0.76	1.90	10.82	256.00	15.50	10.15
Ejelución 3	0.57	0.00	0.57	2.45	21.92	143.00	23.84	169.20	0.76	0.00	0.76	2.46	14.29	0.00	14.29	13.57
Ejelución 4	0.57	0.00	0.57	2.86	21.90	151.00	23.92	180.42	0.76	0.00	0.76	2.82	10.81	255.00	15.47	11.35
Ejelución 5	0.57	0.00	0.57	2.89	21.76	155.00	23.84	288.39	0.76	0.00	0.76	2.58	10.81	266.00	15.67	11.64
Media	0.57	0.00	0.57	2.85	21.23	200.50	23.92	190.58	0.76	0.00	0.76	2.42	12.20	155.40	15.04	11.51

Tabla 2: Resultados obtenidos por el algoritmo BL en el PAR con 20 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T
Ejelución 1	39.66	0.00	0.57	179.71	21.43	94.00	23.95	348.77	0.76	0.00	0.76	152.04	13.83	6.00	14.05	28.88
Ejelución 2	0.57	0.00	0.57	171.87	19.63	96.00	22.21	349.76	0.76	0.00	0.76	167.33	10.81	113.00	14.94	30.86
Ejelución 3	0.57	0.00	0.57	166.51	22.23	78.00	24.32	292.28	0.76	0.00	0.76	164.14	13.83	6.00	14.05	31.16
Ejelución 4	0.57	0.00	0.57	172.49	21.92	87.00	24.25	163.89	0.76	0.00	0.76	161.20	13.83	6.00	14.05	29.67
Ejelución 5	0.57	0.00	0.57	179.71	21.88	97.00	24.48	134.16	0.76	0.00	0.76	141.10	13.83	6.00	14.05	33.47
Media	8.47	0.00	0.57	174.06	21.42	90.40	23.84	257.77	0.76	0.00	0.76	157.16	13.23	27.40	14.23	30.81

Tabla 3: Resultados obtenidos por el algoritmo SA en el PAR con 10 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T
Ejelución 1	0.57	0.00	0.57	221.13	21.98	170.00	24.26	297.10	0.76	0.00	0.76	173.76	10.87	234.00	15.15	41.49
Ejelución 2	0.57	0.00	0.57	219.80	21.89	186.00	24.39	315.58	0.76	0.00	0.76	169.22	14.29	0.00	14.29	45.08
Ejelución 3	0.57	0.00	0.57	219.75	21.90	179.00	24.31	281.15	0.76	0.00	0.76	177.54	14.08	11.00	14.28	45.75
Ejelución 4	0.57	0.00	0.57	229.10	19.63	272.00	23.28	275.46	0.76	0.00	0.76	172.92	14.29	0.00	14.29	57.92
Ejelución 5	0.57	0.00	0.57	225.44	21.50	221.00	24.46	382.79	0.76	0.00	0.76	172.13	14.29	0.00	14.29	45.74
Media	0.57	0.00	0.57	223.04	21.38	205.60	24.14	310.42	0.76	0.00	0.76	173.11	13.56	49.00	14.46	47.20

Tabla 4: Resultados obtenidos por el algoritmo SA en el PAR con 20 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T	Tasa_C	Tasa_inf	Agr,	T
Ejelución 1	0.57	0.00	0.57	171.60	21.43	94.00	23.95	335.60	0.76	0.00	0.76	140.22	13.83	6.00	14.05	92.91
Ejelución 2	0.57	0.00	0.57	182.90	19.63	96.00	22.20	351.39	0.76	0.00	0.76	137.85	13.83	6.00	14.05	77.92
Ejelución 3	0.57	0.00	0.57	164.12	22.86	140.00	26.61	625.01	0.76	0.00	0.76	131.44	13.83	6.00	14.05	96.17
Ejelución 4	0.57	0.00	0.57	168.20	21.49	171.00	26.08	622.46	0.76	0.00	0.76	135.36	13.83	6.00	14.05	84.24
Ejelución 5	0.57	0.00	0.57	171.59	22.01	182.00	26.89	615.38	0.76	0.00	0.76	130.43	13.83	6.00	14.05	95.82
Media	0.57	0.00	0.57	171.68	21.48	136.60	25.15	509.97	0.76	0.00	0.76	135.06	13.83	6.00	14.05	89.41

Tabla 5: Resultados obtenidos por el algoritmo BMB en el PAR con 10 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T
Ejelución 1	0.57	0.00	0.57	230.82	22.43	187.00	24.94	806.14	0.76	0.00	0.76	182.30	14.29	0.00	14.29	103.76
Ejelución 2	0.57	0.00	0.57	235.22	22.15	276.00	25.86	770.85	0.76	0.00	0.76	180.12	14.29	0.00	14.29	98.32
Ejelución 3	0.57	0.00	0.57	237.48	22.59	256.00	26.02	818.17	0.76	0.00	0.76	186.82	14.29	0.00	14.29	99.41
Ejelución 4	0.57	0.00	0.57	231.46	21.83	183.00	24.28	820.18	0.76	0.00	0.76	180.80	14.29	0.00	14.29	105.56
Ejelución 5	0.57	0.00	0.57	237.42	22.60	201.00	25.29	856.11	0.76	0.00	0.76	186.82	14.29	0.00	14.29	117.31
Media	0.57	0.00	0.57	234.48	22.32	220.50	25.28	814.29	0.76	0.00	0.76	183.37	14.29	0.00	14.29	104.87

Tabla 6: Resultados obtenidos por el algoritmo BMB en el PAR con 20 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T
Ejecución 1	0.67	0.00	0.67	27.06	21.14	167.00	171.82	25.62	0.76	0.00	0.76	25.06	13.83	6.00	14.05	84.04
Ejecución 2	0.67	0.00	0.67	26.62	22.24	86.00	24.55	644.13	0.76	0.00	0.76	26.48	13.83	6.00	14.05	90.09
Ejecución 3	0.67	0.00	0.67	29.64	22.39	162.00	26.74	615.09	0.76	0.00	0.76	28.42	13.83	6.00	14.05	80.32
Ejecución 4	0.67	0.00	0.67	29.14	21.93	144.00	25.79	584.98	0.76	0.00	0.76	26.85	13.83	6.00	14.05	79.39
Ejecución 5	1.86	349.00	23.99	30.04	22.29	99.00	24.95	635.15	0.76	0.00	0.76	28.81	13.83	6.00	14.05	101.31
Media	0.91	69.80	5.33	28.50	22.00	131.60	54.77	500.99	0.76	0.00	0.76	27.12	13.83	6.00	14.05	87.03

Tabla 7: Resultados obtenidos por el algoritmo ILS en el PAR con 10 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T
Ejecución 1	0.67	0.00	0.67	30.96	22.01	222.00	24.99	806.16	0.76	0.00	0.76	25.06	14.29	0.00	14.29	107.18
Ejecución 2	0.67	0.00	0.67	33.22	22.36	199.00	25.03	838.21	0.76	0.00	0.76	26.48	14.29	0.00	14.29	104.47
Ejecución 3	0.67	0.00	0.67	27.65	21.93	191.00	24.49	823.25	0.76	0.00	0.76	28.42	14.29	0.00	14.29	103.43
Ejecución 4	0.67	0.00	0.67	32.24	21.51	238.00	24.70	790.53	0.76	0.00	0.76	26.85	14.29	0.00	14.29	111.33
Ejecución 5	0.67	0.00	0.67	36.65	21.53	256.00	24.96	844.69	0.76	0.00	0.76	28.81	14.29	0.00	14.29	111.05
Media	0.67	0.00	0.67	32.14	21.87	221.20	24.83	820.57	0.76	0.00	0.76	27.12	14.29	0.00	14.29	107.49

Tabla 8: Resultados obtenidos por el algoritmo ILS en el PAR con 20 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T
Ejecución 1	1.00	218.00	2.39	10.30	28.99	589.00	44.79	383.19	1.17	170.00	2.42	8.34	15.95	199.00	23.23	36.05
Ejecución 2	1.00	179.00	2.14	10.28	42.20	1432.00	80.63	67.43	1.46	216.00	3.04	8.49	15.45	313.00	26.90	37.13
Ejecución 3	0.85	172.00	1.94	10.29	41.57	1448.00	80.42	73.34	1.47	185.00	2.83	9.66	14.85	269.00	24.69	35.68
Ejecución 4	0.94	216.00	2.31	9.42	42.02	1447.00	80.84	73.37	1.29	185.00	2.64	6.69	17.49	176.00	23.93	39.20
Ejecución 5	0.92	165.00	1.96	9.89	42.62	1424.00	80.53	72.92	1.09	108.00	1.89	8.53	14.29	283.00	24.64	41.39
Media	0.94	190.00	0.67	10.04	39.48	1268.00	73.44	134.05	1.30	172.80	2.56	8.34	15.61	248.00	24.68	37.89

Tabla 9: Resultados obtenidos por el algoritmo ILS-SA en el PAR con 10 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	T
Ejecución 1	0.78	245.00	1.56	13.78	42.06	2888.00	80.80	113.36	1.09	299.00	2.19	12.02	14.12	658.00	26.15	51.82
Ejecución 2	0.81	311.00	1.80	13.89	40.69	2933.00	80.03	105.73	1.15	311.00	2.29	13.19	14.49	419.00	22.15	53.38
Ejecución 3	1.04	379.00	2.24	12.69	41.10	2817.00	78.89	107.36	1.21	313.00	2.36	11.26	15.63	501.00	24.79	55.88
Ejecución 4	0.81	301.00	1.76	12.74	40.47	3002.00	80.74	111.86	1.33	417.00	2.86	11.94	14.72	501.00	23.88	52.00
Ejecución 5	0.79	276.00	1.66	13.24	40.94	2955.00	80.58	110.58	1.1	294.00	2.17	13.42	15.91	524.00	25.48	55.90
Media	0.85	302.40	0.67	13.27	41.05	2919.00	80.21	109.78	1.18	326.80	2.37	12.37	14.97	520.60	24.49	53.80

Tabla 10: Resultados obtenidos por el algoritmo ILS-SA en el PAR con 20 % de restricciones

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>
Ejecución 1	0,67	0,00	0,67	204,29	21,67	69,00	23,53	1907,25	0,76	0,00	0,76	132,57	13,83	6,00	14,05	279,85
Ejecución 2	0,67	0,00	0,67	212,28	22,03	79,00	24,14	1725,83	0,76	0,00	0,76	132,12	13,83	6,00	14,05	288,32
Ejecución 3	0,67	0,00	0,67	218,38	19,60	122,00	22,88	1754,49	0,76	0,00	0,76	130,44	13,83	6,00	14,05	278,83
Ejecución 4	0,67	0,00	0,67	205,21	22,19	61,00	23,82	1632,48	0,76	0,00	0,76	129,51	13,83	6,00	14,05	278,12
Ejecución 5	0,67	0,00	0,67	204,27	21,25	107,00	24,13	1832,01	0,76	0,00	0,76	130,34	13,83	6,00	14,05	277,74
Media	0,67	0,00	0,67	208,89	21,35	87,60	23,70	1770,41	0,76	0,00	0,76	131,00	13,83	6,00	14,05	280,57

**Tabla 11: Resultados obtenidos por el algoritmo FA-LS en el PAR con 10 % de restricciones**

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>
Ejecución 1	0,67	0,00	0,67	204,29	21,78	164,00	23,98	1807,12	0,76	0,00	0,76	132,57	13,83	6,00	14,05	279,85
Ejecución 2	0,67	0,00	0,67	212,28	21,79	171,00	24,09	2017,57	0,76	0,00	0,76	132,12	13,83	6,00	14,05	288,32
Ejecución 3	0,67	0,00	0,67	218,38	21,85	161,00	24,01	2048,20	0,76	0,00	0,76	130,44	13,83	6,00	14,05	278,83
Ejecución 4	0,67	0,00	0,67	205,21	21,91	168,00	24,16	1884,67	0,76	0,00	0,76	129,51	13,83	6,00	14,05	278,12
Ejecución 5	0,67	0,00	0,67	204,27	21,38	158,00	23,90	1832,01	0,76	0,00	0,76	130,34	13,83	6,00	14,05	277,74
Media	0,67	0,00	0,67	208,89	21,74	164,40	24,03	1917,91	0,76	0,00	0,76	131,00	13,83	6,00	14,05	280,57

**Tabla 12: Resultados obtenidos por el algoritmo FA-LS en el PAR con 20 % de restricciones**

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>
BL	0,67	0,00	0,67	1,97	22,36	95,00	24,91	180,23	0,76	0,00	0,76	1,62	11,44	87,40	14,63	8,91
SA	8,47	0,00	0,67	174,06	21,42	90,40	23,84	257,77	0,76	0,00	0,76	157,16	13,23	27,40	14,23	30,81
BMB	0,67	0,00	0,67	171,68	21,48	136,60	25,15	509,97	0,76	0,00	0,76	135,06	13,83	6,00	14,05	89,41
ILS	0,91	69,80	5,33	28,50	22,00	131,60	54,77	500,99	0,76	0,00	0,76	27,12	13,83	6,00	14,05	87,03
ILS-SA	0,94	190,00	0,67	10,04	39,48	1268,00	73,44	134,05	1,30	172,80	2,56	8,34	15,61	248,00	24,68	37,89
FALS	0,67	0,00	0,67	208,89	21,35	87,60	23,70	1770,41	0,76	0,00	0,76	131,00	13,83	6,00	14,05	280,57

**Tabla 13: Resultados globales en el PAR con 10 % de restricciones**

	Iris				Ecoli				Rand				Newthyroid			
	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>	<i>Tasa_C</i>	<i>Tasa_inf</i>	<i>Agr,</i>	<i>T</i>
BL	0,67	0,00	0,67	2,85	21,23	200,60	23,92	190,68	0,76	0,00	0,76	2,42	12,20	155,40	15,04	11,51
SA	0,67	0,00	0,67	223,04	21,38	205,60	24,14	310,42	0,76	0,00	0,76	173,11	13,56	49,00	14,46	47,20
BMB	0,67	0,00	0,67	234,48	22,32	220,60	25,28	814,29	0,76	0,00	0,76	183,37	14,29	0,00	14,29	104,87
ILS	0,67	0,00	0,67	32,14	21,87	221,20	24,83	820,57	0,76	0,00	0,76	27,12	14,29	0,00	14,29	107,49
ILS-SA	0,85	302,40	0,67	13,27	41,05	2919,00	80,21	109,78	1,18	326,80	2,37	12,37	14,97	520,60	24,49	53,80
FALS	0,67	0,00	0,67	208,89	21,74	164,40	24,03	1917,91	0,76	0,00	0,76	131,00	13,83	6,00	14,05	280,57

**Tabla 14: Resultados globales en el PAR con 20 % de restricciones**

De las tablas he suprimido el Greedy, ya que no aporta mucho al análisis por su baja calidad de soluciones en ecoli. Y los genéticos porque eran bastante malos para ser comparados(les he estado mirando estos días pero no he conseguido encontrar nada), al ser estos quitados aun obteniendo muy buenos resultados en el memético tampoco lo he incluido por si un error en los genéticos dan lugar a algún funcionamiento extraño, aunque adelanto que los meméticos han funcionado mejor que cualquier otro que haya implementado, incluyendo el FA.

No he ejecutado todas las semillas del FA original ya que al ejecutarlo una me daban resultados malos tardando más que el FA con LS, pero seguí las trazas del resto de semillas y tenían la misma pinta. Exploraba demasiado y no profundizaba en soluciones decentes. No se si esto es debido a que la representación y movimiento de las luciérnagas

sea una limitación a la funcionalidad del algoritmo. Seguramente con un estudio mejor de los parámetros y adaptándolos al problema, funcionaría algo mejor el FA original.

Para mejorar la calidad de la solución le incorporé una LS suave como en el memético, pero no fue suficiente. Tras ver que no mejoraba lo suficiente le añadí una LS como la de la P1, y la cosa mejoró bastante. Con este FA-LS se obtienen soluciones de mejor calidad que la propia LS y que el SA. La alta exploración de soluciones del FA funciona bien con la explotación de la LS. Esta LS se hace cada 10 bucles for(i), como las 10 generaciones del memético.

Como ya he dicho anteriormente, como conclusión final pienso que se podría aumentar la efectividad del FA haciendo un buen estudio de los parámetros en vez de usar los por defecto. Además, en uno de los paper que he leído sobre un problema de restricciones[4](parecido a los de la asignatura TSI) usa un esquema de enfriamiento como el que usamos en el SA en la P2 para modificar la componente aleatoria del movimiento. Podría estudiarse si eso ayuda al algoritmo.

## 5. Manual de uso

Si se desea ejecutar:

`$main.py < dataset > < n_restricciones > < algoritmo >`

algoritmos:

- BL - bl
- SA - sa
- BMB - bmb
- ILS - ils ls
- ILS/SA - ils sa
- FA - fa

Si se ejecuta un ILS:

`$main.py < dataset > < n_restricciones > ils [ls|sa]`

Si se ejecuta FA, para los parámetros de la LS:

`$main.py < dataset > < n_restricciones > fa 1|0,1 better|normal`

Para ejecutar con una semilla concreta, dentro de *main.py* hay un array con las semillas usadas según el orden de aparición en las tablas. Basta con cambiar la asignación de la variable *seed* a *seeds[i]* correspondiente a la Ejecución *i*.

## Referencias

- [1] Xin-She Yang. Firefly algorithms for multimodal optimization. *Conference Paper*, March 2010.
- [2] Monika Bajaj Hema Banati. Performance analysis of firefly algorithm for data clustering. *Conference Paper*, January 2013.
- [3] Hailun Xie, Li Zhang, Chee Peng Lim, Yonghong Yu, Chengyu Liu, Han Liu, and Julie Walters. Improving k-means clustering with enhanced firefly algorithms. *Conference Paper*, November 2019.
- [4] Nebojsa Bacanin Ivona Brajevic, Milan Tuba. Firefly algorithm with a feasibility-based rules for constrained optimization. *Conference Paper*, September 2012.
- [5] Daniel Molina, Javier Poyatos, Javier Del Ser, Salvador García, Amir Hussain, and Francisco Herrera. Comprehensive taxonomies of nature- and bio-inspired optimization: Inspiration versus algorithmic behavior, critical analysis and recommendations. *Computer Science Paper*, page 76, Feb 2020.