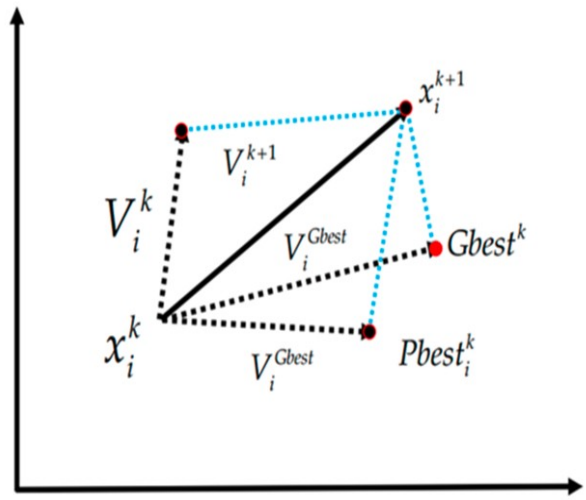
Algoritmos Evolutivos Tarea 4.

Particle Swarm Optimization (PSO)

Introducción.

El algoritmo PSO original fue inspirado por el comportamiento social de algunos organismos biológicos, específicamente la habilidad de algunos grupos de especies de cooperar haciendo formaciones especiales en un área determinada, por ejemplo, una parvada de aves dirigiéndose a una fuente de alimento.



En las implementaciones mas comunes del PSO, las partículas se desplazan por el espacio de búsqueda siguiendo una atracción hacia la mejor solución que individualmente han encontrado, y una atracción hacia la mejor solución que cualquier partícula en su *vecindad* ha encontrado. Se define *vecindad* como el conjunto de partículas entre las cuales existe comunicación.

Una partícula i esta compuesta por tres vectores: su posición $x_i=(x_{i1},x_{i2},\dots,x_{id})$ en el espacio de dimensión d, su velocidad $v_i=(v_{i1},v_{i2},\dots,v_{id})$ y la mejor posición que individualmente a visitado $p_i=(p_{i1},p_{i2},\dots,p_{id})$. Tanto la posición como la velocidad de cada partícula se inicializa de manera aleatoria.

Estas partículas se desplazan a lo largo del espacio de búsqueda por medio de unas ecuaciones de actualización bastante simples. El algoritmo PSO clásico actualiza el enjambre entero en cada paso de tiempo actualizando la velocidad y posición de cada partícula para cada dimensión usando las siguientes reglas:

$$v_i^j = v_i^j + c\,\varepsilon_1(p_i^j - x_i^j) + c\,\varepsilon_2(l_{best}^j - x_i^j) \quad (1)$$

$$x_i^j = x_i^j + v_i^j \quad (2)$$

En las ecuaciones original, c = 2, ε_1, ε_2 son valores aleatorios que se generan en cada iteración, y l_{best} es la mejor posición encontrada en la vecindad de la partícula. El proceso de actualización se resume en el siguiente algoritmo:

Algorithm 1 The PSO update process

```
for each time step t do
  for each particle i in the swarm do
    update position x⃗_t using eqs 1 & 2
    calculate particle fitness f(x⃗_t)
    update p⃗_i, p⃗_g
  end for
end for
```

Implementación.

En esta práctica se implementa una versión del algoritmo PSO que utiliza un factor de constricción en la actualización de las velocidades, además de constantes de aceleración fijas. El algoritmo fue probado con dos de las topologías mas conocidas, a saber, la topología de conectividad completa y la topología de anillo. Además, se propone una tercera topología y se comparan resultados utilizando un benchmark de funciones unimodales y multimodales:

Sphere	$f_{sph}(\vec{\mathbf{x}}) = \sum_{i=1}^{N} x_i^2$	$[-100, 100]^{30}$
Elliptic	$f_{ell}(\vec{\mathbf{x}}) = \sum_{i=1}^{N} \left(10^6\right)^{\left(\frac{i-1}{N-1}\right)} x_i^2$	$[-100, 100]^{30}$
Schwefel 1.2	$f_{sch}(\vec{\mathbf{x}}) = \sum_{i=1}^{N} \left(\sum_{j=1}^{i} x_j\right)^2$	$[-100, 100]^{30}$
Ackley	$f_{ack}(\vec{\mathbf{x}}) = 20 + e - 20e^{\left(-0.2\sqrt{\left(\frac{1}{N}\sum_{i=1}^{N} x_i^2\right)}\right)} - e^{\left(\frac{1}{N}\sum_{i=1}^{N}\cos(2\pi x_i)\right)}$	$[-32, 32]^{30}$
Rastrigin	$f_{ras}(\vec{\mathbf{x}}) = 10N + \sum_{i=1}^{N}(x_i^2 - 10\cos(2\pi x_i))$	$[-5.12, 5.12]^{30}$
Griewank	$f_{grw}(\vec{\mathbf{x}}) = \sum_{i=1}^{N} x_i^2/4000 - \prod_{i=1}^{N}\cos(x_i/\sqrt{i}) + 1$	$[-600, 600]^{30}$
Rosenbrock	$f_{ros}(\vec{\mathbf{x}}) = \sum_{i=1}^{N-1}(100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2)$	$[-100, 100]^{30}$
Weierstrass	$f_{wrs}(\vec{\mathbf{x}}) = \sum_{i=1}^{N}\left(\sum_{k=0}^{k_{max}}(a^k\cos(2\pi b^k(x_i + 0.5)))\right) - N\sum_{k=0}^{k_{max}}(a^k\cos(\pi b^k)$	$[-0.5, 0.5]^{30}$
Schaffer	$f_{scf}(\vec{\mathbf{x}}) = \sum_{i=1}^{N} F(x_i, x_{i+1}); \quad x_{N+1} = x_1$ where, $F(x, y) = 0.5 + \frac{sin^2(\sqrt{x^2+y^2}) - 0.5}{(1+0.001(x^2+y^2))^2}$	$[-0.5, 0.5]^{30}$
Salomon	$f_{sal}(\vec{\mathbf{x}}) = 1 - \cos(2\pi\|\|\mathbf{x}\|\|) + 0.1\|\|\mathbf{x}\|\|$ where, $\|\|\mathbf{x}\|\| = \sqrt{\left(\sum_{i=1}^{N} x_i^2\right)}$	$[-100, 100]^{30}$

Fórmula de actualización de la velocidad de las partículas para el método *Constriction Factor*:

$$v_i^j = \chi\left(v_i^j + c_1\varepsilon_1\left(p_{ibest}^j - x_i^j\right) + c_2\varepsilon_2\left(l_{best}^j - x_i^j\right)\right)$$

Aquí, p_{ibest} es la mejor posición visitada por la i-ésima partícula, l_{best} es la mejor posición que cada vecindad ha visitado (definido de acuerdo a cada topología), ε_1, ε_2 son valores aleatorios uniformemente distribuidos en [0,1] y χ es el factor de constricción que se define como:

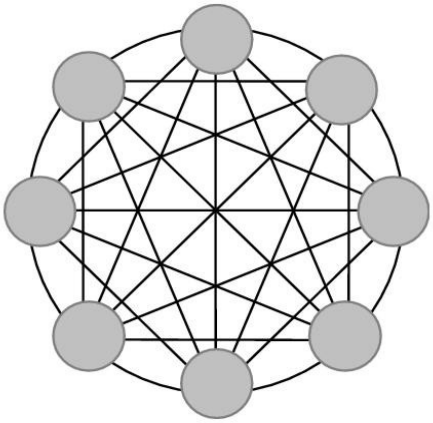
$$\chi = \frac{2}{\left|2 - \varphi - \sqrt{\varphi^2 - 4\varphi}\right|} \quad \text{con} \quad \varphi = c_1 + c_2$$

donde $c_1 = c_2 = 2.05$; de modo que $\varphi = 4.1$ y $\chi = 0.72984$. La posición de cada partícula se actualiza de la manera clásica:

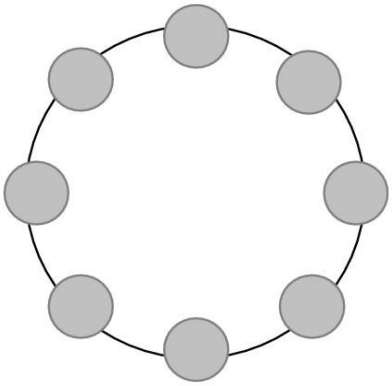
$$x_i^j = x_i^j + v_i^j$$

Topologías.

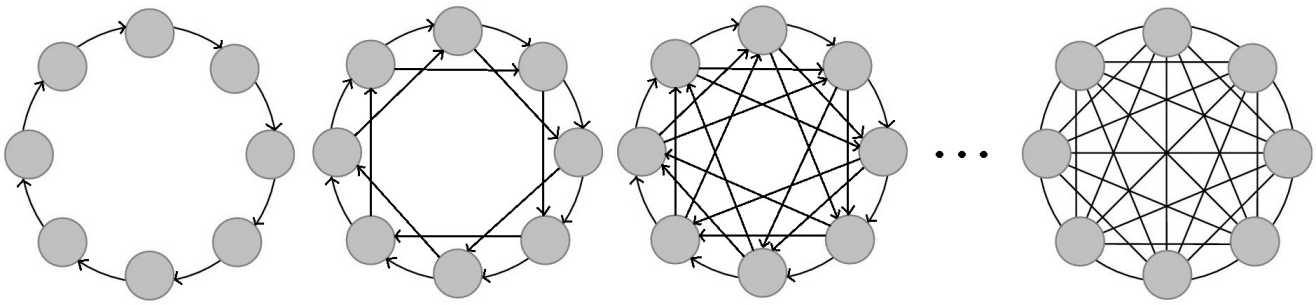
1) Conectividad completa. La vecindad de cada partícula equivale a la población entera, de modo que el mejor local l_{best} equivale al mejor global g_{best}.



2) Conectividad de anillo. Cada partícula tiene exactamente dos vecinos y ella misma pertenece a dos vecindades distintas.



3) Topología propuesta. Consiste en una conectividad dinámica, en la cual cada partícula empieza teniendo solamente un vecino y ella misma pertenece a una sola vecindad, y conforme se incrementan las iteraciones, se va agregando una partícula a cada vecindad hasta terminar con una vecindad completa. La idea es que se se formen al principio muchos “clusters” de partículas donde haya un mejor local, y que conforme transcurra el tiempo los clusters se vayan agrupando poco hasta tener un solo enjambre.



Resultados.

La Tabla 1 muestra los resultados que se obtuvieron con 50 ejecuciones del algoritmo utilizando la topología de conectividad completa, el criterio de paro se fijó a 1000 x *d* evaluaciones de función, el *S_ratio* indica el porcentaje de esas ejecuciones donde se alcanzó un valor de fitness por debajo de $1x10^{8}$.

Table 1. Complete Topology							
Function	**Dim**	**Fmin**	**Fmax**	**Favg**	**Fmedian**	**FstdDev**	**S_ratio**
Sphere	30	1.60E-173	1.88E-153	4.25E-155	2.96E-164	2.63E-154	100
Sphere	50	1.31E-104	3.31E-087	6.79E-089	2.67E-096	4.63E-088	100
Elliptic	30	2.29E-168	5.15E-138	1.03E-139	2.71E-158	7.21E-139	100
Elliptic	50	3.85E-104	3.50E-080	8.23E-082	2.06E-091	4.96E-081	100
Schwefel	30	2.65E-026	5.82E-022	5.76E-023	1.03E-023	1.23E-022	100
Schwefel	50	2.00E-011	5.29E-008	2.44E-009	5.00E-010	7.56E-009	96
Ackley	30	1.47E-014	8.13E+000	2.22E+000	1.90E+000	1.43E+000	10
Ackley	50	2.43E+000	1.51E+001	6.66E+000	6.69E+000	2.70E+000	0
Rastrigin	30	4.28E+001	1.16E+002	7.56E+001	7.26E+001	1.69E+001	0
Rastrigin	50	1.11E+002	2.61E+002	1.79E+002	1.72E+002	3.64E+001	0
Griewank	30	0.00E+000	3.16E-001	3.89E-002	1.72E-002	6.27E-002	26
Griewank	50	1.78E-015	7.58E+000	3.58E-001	1.01E-001	1.08E+000	2
Rosenbrock	30	8.85E-005	4.58E+002	3.95E+001	4.05E+000	9.83E+001	0
Rosenbrock	50	1.32E-006	1.01E+002	1.70E+001	4.04E+000	2.60E+001	0
Weierstrass	30	2.47E+000	1.53E+001	6.51E+000	6.18E+000	2.38E+000	0
Weierstrass	50	1.31E+001	2.73E+001	2.13E+001	2.16E+001	3.28E+000	0
Schaffer	30	0.00E+000	3.89E-002	1.17E-003	0.00E+000	6.03E-003	96
Schaffer	50	0.00E+000	1.94E-002	7.77E-004	8.88E-016	3.81E-003	96
Salomon	30	4.00E-001	2.70E+000	6.68E-001	6.00E-001	3.62E-001	0
Salomon	50	7.00E-001	7.30E+000	1.85E+000	1.40E+000	1.26E+000	0

La Tabla 2 muestra los resultados obtenidos con la topología de anillo:

Table 2. Ring Topology							
Function	**Dim**	**Fmin**	**Fmax**	**Favg**	**Fmedian**	**FstdDev**	**S_ratio**
Sphere	30	1.02E-078	3.06E-075	2.78E-076	7.00E-077	4.94E-076	100
Sphere	50	1.97E-075	5.20E-072	7.20E-073	2.45E-073	1.16E-072	100
Elliptic	30	4.98E-075	1.68E-071	1.72E-072	4.26E-073	3.24E-072	100
Elliptic	50	2.34E-071	6.87E-069	8.35E-070	3.66E-070	1.32E-069	100
Schwefel	30	1.14E-010	3.83E-008	1.07E-008	5.79E-009	1.10E-008	62
Schwefel	50	7.80E-001	6.94E+001	5.87E+000	2.77E+000	1.07E+001	0
Ackley	30	4.00E-015	7.55E-015	7.48E-015	7.55E-015	4.97E-016	100
Ackley	50	7.55E-015	1.03E+000	4.11E-002	1.47E-014	2.01E-001	96
Rastrigin	30	3.98E+001	1.12E+002	7.14E+001	7.11E+001	1.56E+001	0
Rastrigin	50	1.15E+002	2.36E+002	1.74E+002	1.76E+002	2.77E+001	0
Griewank	30	0.00E+000	1.72E-002	1.63E-003	0.00E+000	4.02E-003	84
Griewank	50	0.00E+000	9.86E-003	5.42E-004	0.00E+000	2.16E-003	94
Rosenbrock	30	2.79E-002	6.05E+002	4.36E+001	1.10E+001	1.03E+002	0
Rosenbrock	50	3.43E-001	1.15E+003	1.28E+002	3.56E+001	2.41E+002	0
Weierstrass	30	1.42E+000	1.47E+001	8.42E+000	8.67E+000	3.33E+000	0
Weierstrass	50	1.69E+001	3.62E+001	2.61E+001	2.64E+001	5.24E+000	0
Schaffer	30	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100
Schaffer	50	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100
Salomon	30	2.00E-001	3.00E-001	2.46E-001	2.00E-001	4.98E-002	0
Salomon	50	3.00E-001	5.00E-001	3.60E-001	4.00E-001	5.66E-002	0

Por último, la Tabla 3 muestra lo que se obtuvo al utilizar la topología propuesta:

Table 3. Proposed Topology								
Function	**Dim**	**Fmin**	**Fmax**	**Favg**	**Fmedian**	**FstdDev**	**S_ratio**	**Iterations**
Sphere	30	1.24E-168	3.23E-150	1.12E-151	6.53E-159	5.56E-151	100	10000
Sphere	50	2.08E-166	4.02E-150	9.82E-152	3.09E-157	5.63E-151	100	16666
Elliptic	30	1.74E-162	1.22E-145	2.59E-147	1.56E-154	1.70E-146	100	10000
Elliptic	50	1.38E-164	3.51E-143	7.20E-145	3.85E-151	4.91E-144	100	16666
Schwefel	30	4.00E-042	1.02E-034	5.56E-036	6.71E-038	1.91E-035	100	10000
Schwefel	50	3.90E-044	9.84E-036	4.66E-037	1.20E-039	1.65E-036	100	16666
Ackley	30	4.00E-015	7.55E-015	6.63E-015	7.55E-015	1.56E-015	100	10000
Ackley	50	4.44E-016	4.44E-016	4.44E-016	4.44E-016	0.00E+000	100	16666
Rastrigin	30	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	10000
Rastrigin	50	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	16666
Griewank	30	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	10000
Griewank	50	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	16666
Rosenbrock	30	9.92E-009	7.58E-008	1.24E-008	9.99E-009	1.16E-008	94	44857
Rosenbrock	50	9.92E-009	1.52E-008	1.01E-008	9.99E-009	7.28E-010	98	78346
Weierstrass	30	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	10000
Weierstrass	50	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	16666
Schaffer	30	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	10000
Schaffer	50	0.00E+000	0.00E+000	0.00E+000	0.00E+000	0.00E+000	100	16666
Salomon	30	5.75E-054	9.99E-002	7.33E-003	8.62E-013	2.51E-002	70	10000
Salomon	50	8.78E-049	9.99E-002	2.87E-003	4.33E-023	1.46E-002	90	16666

Conclusiones.

Primero, si comparamos el PSO con las dos primeras topologías utilizadas vemos que para funciones como la esfera, en las que se necesita mucha intensificación, la topología del anillo no permite alcanzar mínimos tan buenos como la topología completa, sin embargo en varios casos la topología de anillo resulta ser mas eficiente, tal es el caso de funciones multimodales como la Ackley y la Griewank.

En cuanto a la topología propuesta, se observan resultados bastante buenos para casi todas la funciones del benchmark, como se esperaba, el algoritmo resulta ser muy bueno para funciones multimodales, debido a la estructura diversificada de la topología durante las primeras iteraciones, pero también resulta ser bastante bueno con funciones unimodales, debido a la estructura unificada de la topología que se forma en las ultimas iteraciones.

Sin embargo, en funciones como la Rosenbrock, el algoritmo presenta muchos problemas para tratar de minimizar la función dentro del numero de evaluaciones requeridas.

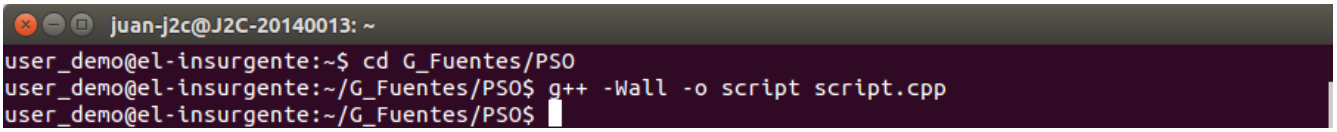
Compilación/Ejecución.

Los programas implementados incluyen un makefile para compilarse, que soporta los comandos *make*, *make run* y *make clean*, así como un script de ejecución en el cluster para correr todas las instancias, los resultados se generan en un archivo para cada ejecución.

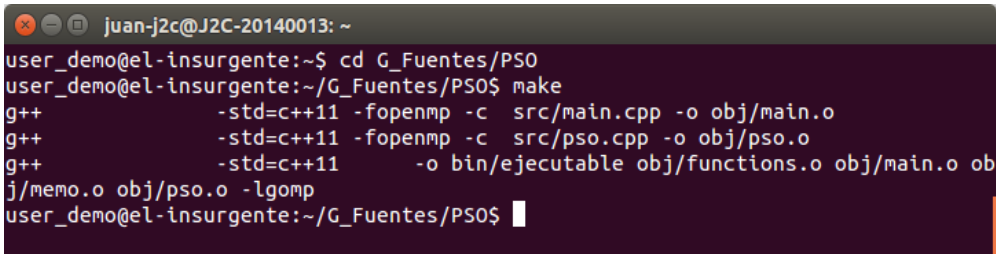
Los programas reciben tres parámetros:

1. Tamaño de la población.
2. Dimensión de las variables (se utilizó 30 y 50).
3. Referencia a la función objetivo a evaluar:
   - [0] = Sphere;
   - [1] = Elliptic;
   - [2] = Schwefel;
   - [3] = Ackley;
   - [4] = Rastrigin;
   - [5] = Griewank;
   - [6] = Rosenbrock;
   - [7] = Weierstrass;
   - [8] = Schaffer;
   - [9] = Salomon;
4. Referencia a la topología que se va a utilizar.
   - [1] = Conectividad completa
   - [2] = Conectividad de anillo
   - [3] = Conectividad dinámica

Compilación del script:



Compilación del programa:



Ejecución del script: