

Introducción a la Computación Evolutiva

Dr. Carlos A. Coello Coello

Departamento de Computación

CINVESTAV-IPN

Av. IPN No. 2508

Col. San Pedro Zacatenco

México, D.F. 07300

email: ccoello@cs.cinvestav.mx

[http: //delta.cs.cinvestav.mx/~ccoello](http://delta.cs.cinvestav.mx/~ccoello)

AGs Paralelos

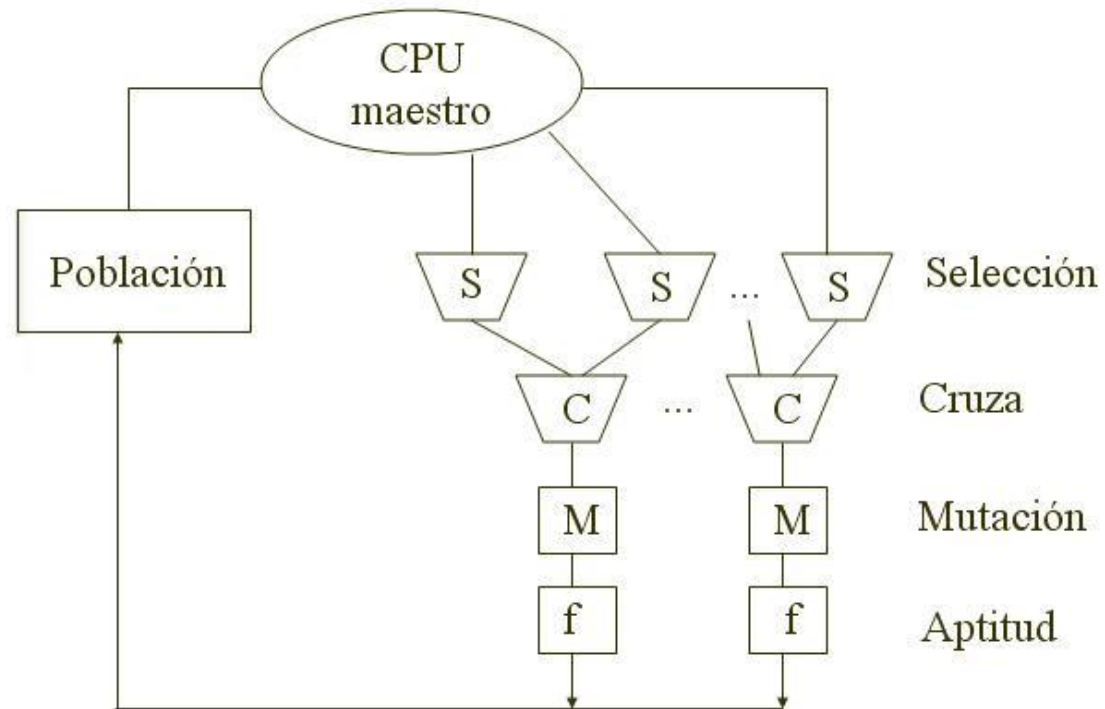
El método más simple de paralelizar un AG es la llamada **paralelización global**.

En este caso, sólo hay una población, como en el AG convencional, pero la evaluación de los individuos y los operadores genéticos se paralelizan de forma **explícita**.

AGs Paralelos

Puesto que sólo hay una población, la selección considera a todos los individuos y cada individuo tiene oportunidad de aparearse con cualquier otro (o sea, hay **apareamiento aleatorio**). Por lo tanto, el comportamiento del AG simple permanece sin cambios.

AGs Paralelos



AGs Paralelos

La paralelización global es un método relativamente fácil de implementar y puede obtenerse un incremento significativo de velocidad si los costos de comunicación no dominan los costos de procesamiento.

Una observación importante es que no debe confundirse el concepto de paralelismo implícito de un AG (visto antes en clase) con el de **paralelismo explícito**.

AGs Paralelos

A la paralelización global también se le conoce como **AG panmítico**, pues se cuenta con un solo depósito de material genético (*gene pool*), o sea con una sola población.

AGs Paralelos

Los AGs panmíticos son útiles cuando el costo de evaluar la función de aptitud es elevado (p.ej., una simulación).

En el AG panmítico no se requieren nuevos operadores ni nuevos parámetros y la solución encontrada será la misma que la producida con un AG convencional (o sea, serial).

AGs Paralelos

Es importante hacer notar que aunque el paralelismo global normalmente es **síncrono** (o sea, que el programa se detiene y espera a recibir los valores de aptitud de toda la población antes de proceder a producir la siguiente generación), puede implementarse también de forma **asíncrona**, aunque en ese caso, su funcionamiento ya no resultará equivalente al de un AG convencional.

AGs Paralelos

Además de paralelizarse la evaluación de la función de aptitud, en el paralelismo global es posible incluir también los operadores genéticos, pero dada la simplicidad de éstos, no suelen paralelizarse, pues los costos de comunicación disiparían cualquier mejora en el desempeño del programa.

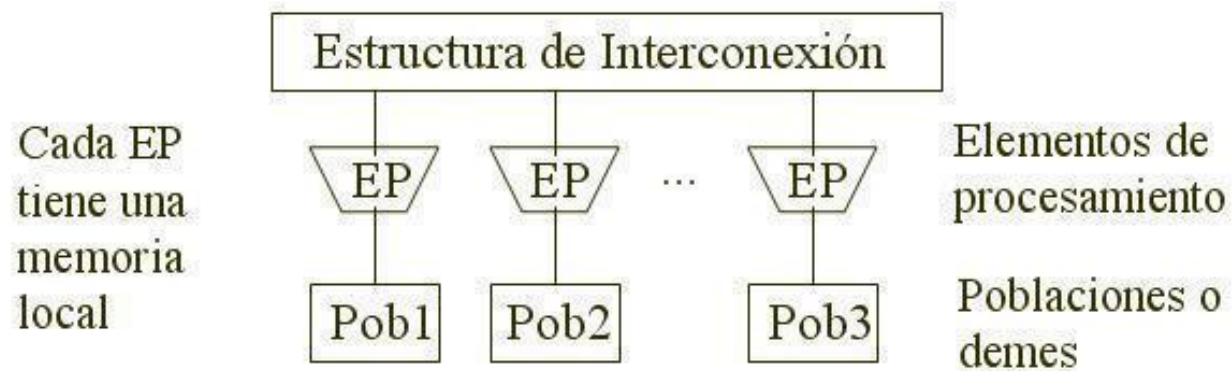
AGs Paralelos

Una idea más sofisticada es usar los llamados **AGs de grano grueso**. En este caso, la población del AG se divide en múltiples subpoblaciones o **demes** que evolucionan de manera aislada la mayor parte del tiempo, aunque intercambian individuos ocasionalmente.

AGs Paralelos

A este intercambio de individuos se le llama **migración**, y se considera como un nuevo operador genético. Además de requerirse parámetros adicionales en este caso, el comportamiento de un AG de grano grueso es diferente del de un AG convencional.

AGs Paralelos



AGs Paralelos

A los **AGs de grano grueso** se les suele llamar también **AGs distribuidos**, porque suelen implementarse en computadoras MIMD con memoria distribuida. Asimismo, algunos autores los llaman también **AGs de isla**, haciendo alusión a un modelo poblacional usado en genética en el cual se consideran **demes** relativamente aislados. A este modelo se le conoce como **modelo de isla**.

AGs Paralelos

Los AGs de grano grueso son muy populares debido a varias razones:

- Son una extensión muy simple de los AGs seriales. Simplemente se toman unos cuantos AGs convencionales (seriales), se corre cada uno de ellos en un procesador diferente y, a ciertos intervalos de tiempo, se intercambian unos cuantos individuos entre ellos.

AGs Paralelos

- Aunque no se tenga acceso a una arquitectura paralela, puede implementarse un AG de grano grueso a través de una simulación efectuada en una red de estaciones de trabajo, o incluso en una computadora con un solo procesador haciendo la simulación mediante software (usando por ejemplo MPI o PVM).

AGs Paralelos

- Se requiere relativamente de poco esfuerzo para convertir un AG serial en un AG de grano grueso. La mayor parte de la programación permanece igual, y sólo se requieren ciertas rutinas adicionales para implementar la migración.

AGs Paralelos

Los parámetros que requieren los AGs de grano grueso son:

- Número de **demes** (esto puede estar determinado por el hardware disponible).
- Tamaño de cada **deme**.
- Estructura de la interconexión (o sea, la **topología**).

AGs Paralelos

Parámetros (continúa):

- Intervalo de migración
- Tasa de migración
- Radio de selección
- Radio de migración

AGs Paralelos

De entre estos parámetros, algunos como la **topología**, juegan un papel preponderante en el desempeño del AG. La topología determina qué tan rápido (o qué tan lentamente) se disemina una buena solución hacia los otros **demes**.

AGs Paralelos

Si se usa una topología dispersamente conectada (con un diámetro grande), las soluciones se diseminarán más lentamente y los **demes** estarán más aislados entre sí, permitiendo la aparición de soluciones diferentes, favoreciendo probablemente la diversidad.

AGs Paralelos

La topología juega también un papel preponderante en el costo de las migraciones.

Por ejemplo, una topología densamente conectada puede promover una mejor mezcla de individuos, pero a un costo computacional más alto.

AGs Paralelos

Se sabe, por ejemplo, que una topología densa tiende a encontrar soluciones globales con un menor número de evaluaciones de la función de aptitud que si se usa una topología dispersa.

AGs Paralelos

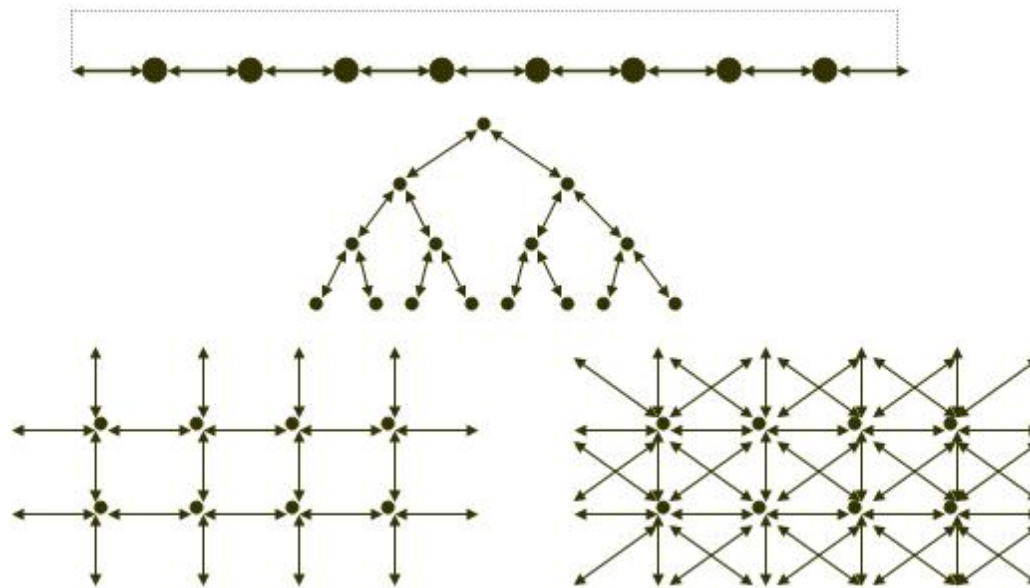
También es posible usar topologías **dinámicas**, en las que los **demes** no están limitados a poder comunicarse sólo con un cierto conjunto predefinido de **demes**, sino que envía sus migrantes a aquellos **demes** que satisfacen ciertos criterios.

AGs Paralelos

La idea de este esquema es que puedan identificarse los **demes** donde los migrantes tienen mayores probabilidades de producir algún efecto. Usualmente, se usa la diversidad como el criterio principal para definir qué tan adecuado es un cierto **deme**.

AGs Paralelos

Algunas topologías posibles son las siguientes:



AGs Paralelos

Es importante mantener en mente la idea de que una topología es una estructura lógica que puede diferir de la estructura de hardware disponible. Es decir, la topología de un AG paralelo no necesariamente debe coincidir con la de nuestra computadora. El problema de hacer esto, sin embargo, es que los costos de comunicación resultantes pueden ser muy elevados.

AGs Paralelos

Relacionado con las topologías se encuentra el concepto de **vecindario**.

El vecindario se refiere al área dentro de la cual puede moverse un migrante de un cierto **deme**.

AGs Paralelos

Asociado al vecindario se encuentra el concepto de **radio de selección**, que se refiere a la cantidad de vecinos entre los cuales se puede efectuar la selección.

Es común usar un radio de selección de cero, o sea, efectuar la selección sólo dentro del mismo **deme**, aunque cualquier otro valor es válido.

AGs Paralelos

Es común usar vecindarios compactos en computación evolutiva, motivados por el hecho de que en la naturaleza, las poblaciones están limitadas geográficamente.

En AGs paralelos, es fácil definir vecindarios compactos, y de ahí que sean tan populares en computación evolutiva.

AGs Paralelos

Otra forma de paralelizar un AG es usando un esquema de **grano fino**.

En este caso, la población de un AG se divide en un gran número de subpoblaciones muy pequeñas. De hecho, el caso ideal sería tener sólo un individuo por cada unidad de procesamiento disponible.

Este modelo es adecuado para arquitecturas masivas en paralelo, aunque puede implementarse en cualquier tipo de multiprocesador.

AGs Paralelos

El problema del paralelismo de grano fino es que el costo de comunicación entre los procesadores puede hacer que el desempeño del algoritmo se degrade con relativa facilidad.

Es común implementar este tipo de paralelismo colocando los individuos en una malla bidimensional, debido a que ésta es la topología usada en hardware para muchas arquitecturas masivas en paralelo.

AGs Paralelos

Resulta difícil comparar de manera justa a un AG paralelo de grano fino con uno de grano grueso, y los pocos estudios al respecto suelen enfatizar sólo una cierta métrica (por ejemplo, la calidad de las soluciones encontradas).

De tal forma, no hay un claro ganador entre estos 2 esquemas.

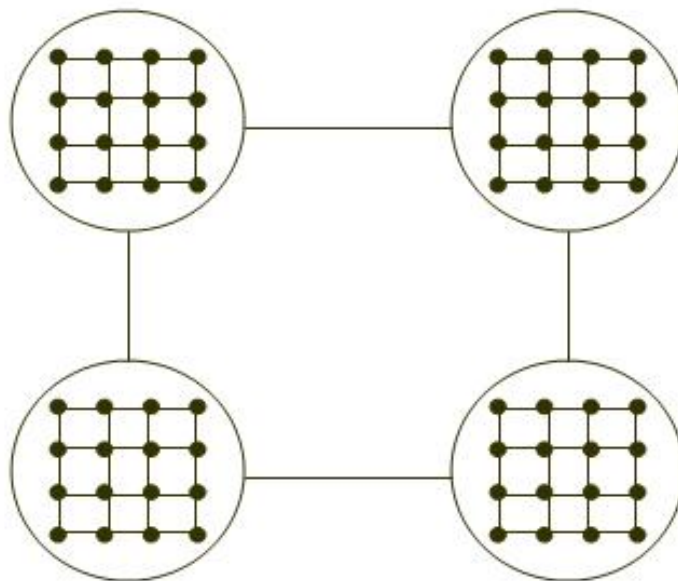
AGs Paralelos

Otra posibilidad para implementar un AG paralelo es combinar los esquemas descritos anteriormente. Debe cuidarse, sin embargo, de que el esquema resultante no sea más complejo que los esquemas originales.

A continuación veremos varios híbridos posibles.

AGs Paralelos

Un posible híbrido consiste en usar un AG de grano fino a bajo nivel y otro de grano grueso a alto nivel.



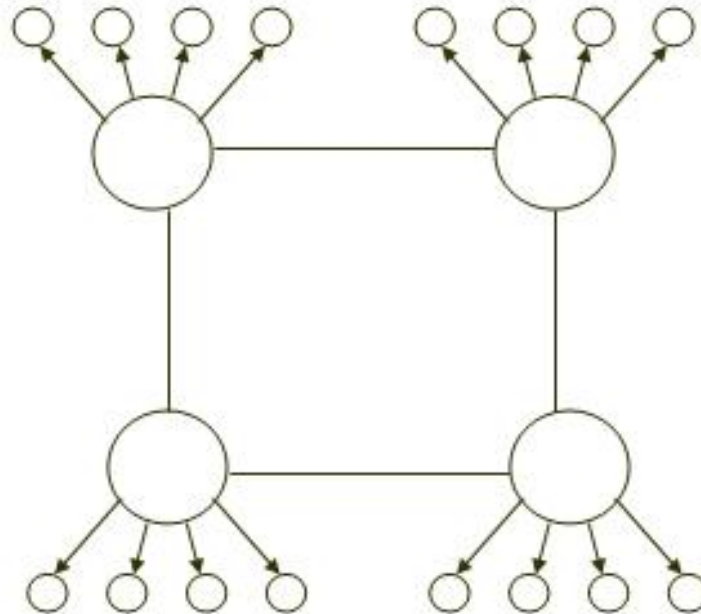
AGs Paralelos

Un ejemplo de este tipo de híbrido es el AG propuesto por Gruau (1994), en el cual la población de cada **deme** se coloca en una malla bidimensional y los **demes** se conectan entre sí en forma de toroide bidimensional.

La migración entre los **demes** vecinos ocurre a intervalos regulares.

AGs Paralelos

Otro posible esquema híbrido consiste en usar una forma de paralelización global en cada uno de los **demes** de un AG de grano grueso.

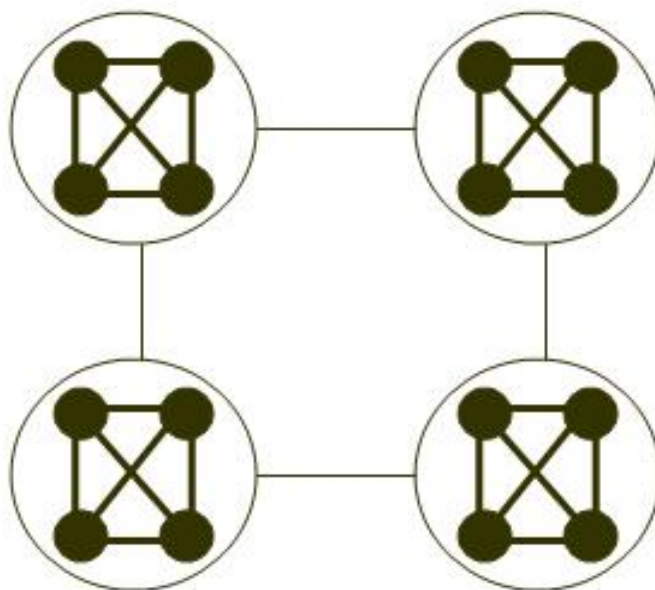


AGs Paralelos

En este caso, la migración ocurre entre los **demes** de manera similar a un AG de grano grueso, pero la evaluación de los individuos se maneja en paralelo. Esta técnica no introduce nuevos problemas analíticos, y puede ser muy útil cuando se trabaja con aplicaciones en las cuales la mayor parte del tiempo de procesamiento lo consume la evaluación de la función de aptitud.

AGs Paralelos

Un tercer método híbrido podría consistir en usar un AG de grano grueso tanto a bajo como a alto nivel.



AGs Paralelos

En este caso, la idea es forzar el mezclado panmítico a bajo nivel usando una alta tasa de migración y una topología densa, y usar una baja tasa de migración a alto nivel. Este híbrido sería equivalente en complejidad a un AG de grano grueso, si consideramos a los grupos de subpoblaciones panmíticas como un solo **deme**. Según Cantú Paz (1997), este esquema nunca ha sido implementado.

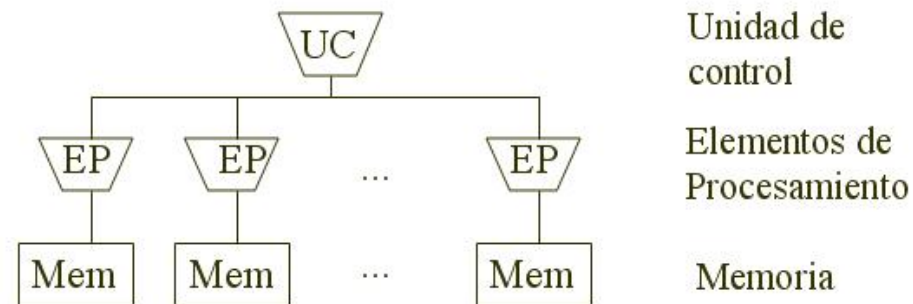
AGs Paralelos

Otra forma de hablar sobre AGs paralelos, es desde la perspectiva del tipo de arquitectura computacional a utilizarse. Desde este punto de vista, podemos hablar fundamentalmente de usar:

- 1) SIMD
- 2) MIMD

AGs Paralelos

Un ejemplo de arquitectura SIMD es el siguiente:



en este caso, cada EP tiene su propia memoria y controla su propio espacio de direccionamiento, aunque también podría haber una sola memoria global compartida por todos los EPs.

AGs Paralelos

- Las arquitecturas SIMD normalmente tienen una forma de malla (*mesh*) o de toroide.
- Esta arquitectura (SIMD) suele usarse para AGs de grano fino (o sea, para **demes** de tamaño reducido: 12 a 15 individuos cada uno).
- Sin embargo, puede implementarse un AG de grano grueso con la misma arquitectura si se usan tamaños mayores de **demes** (p.ej. 50 ó 100 individuos).

AGs Paralelos

- El grado de conectividad es mucho más importante que la estructura de las interconexiones en una arquitectura SIMD. Se sabe que un grado de conectividad de alrededor de 6 es razonablemente bueno.
- El uso de una arquitectura SIMD está asociado con el uso de migración y dicho operador puede complicarse bastante, dependiendo del grado de conectividad de la arquitectura

AGs Paralelos

Dos son los parámetros principales relacionados con la **migración**:

- 1) **Vecindario de migración**: hacia qué demes podemos migrar un individuo.
- 2) **Probabilidad de migración**: ¿cuál es la probabilidad de aceptar a un migrante en un cierto **deme** (suelen usarse valores altos, p.ej. 0.8)?

AGs Paralelos

Los puntos importantes relacionados con la migración son 2:

- 1) ¿A quién importar en una población?
- 2) ¿A quién reemplazar en una población?

AGs Paralelos

Existen varios criterios para llevar a cabo estas 2 operaciones:

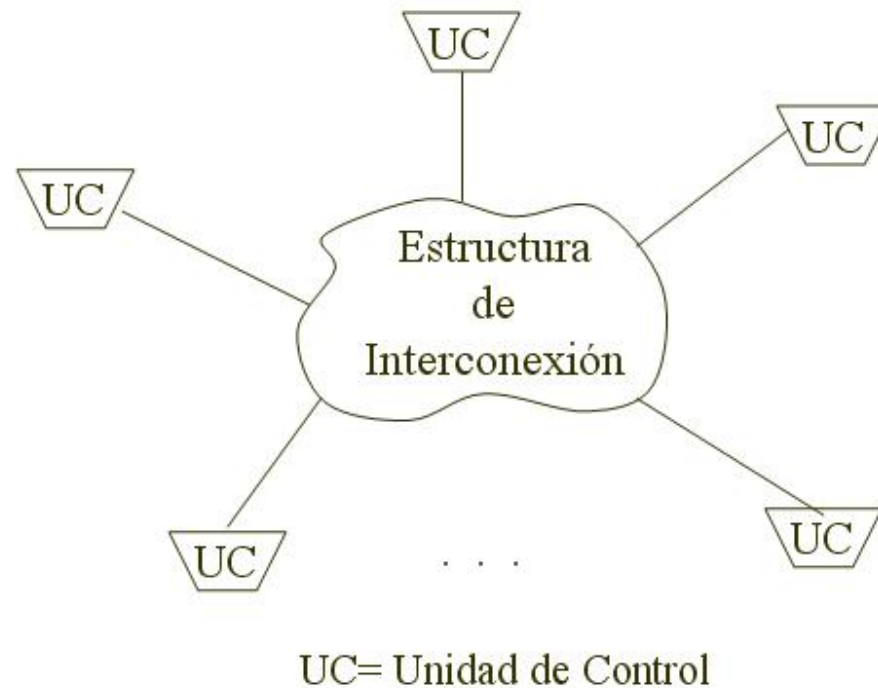
- Importar al azar y reemplazar al azar
- Importar al azar y reemplazar al peor individuo en el **deme**
- Importar el mejor y reemplazar al azar
- Importar el mejor y reemplazar el peor

AGs Paralelos

- Se sabe que la política de reemplazo no es muy importante (no parece tener un efecto significativo en el desempeño de un AG paralelo).
- Sin embargo, el criterio de importación sí es importante (importar al mejor parece funcionar bien).

AGs Paralelos

Por otro lado, tenemos las arquitecturas MIMD:



AGs Paralelos

Las arquitecturas MIMD pueden ser de 2 tipos:

- 1) **Descentralizadas:** tienen poca o ninguna memoria global.
- 2) **Centralizadas:** cuentan con una memoria global compartida.

Las arquitecturas MIMD suelen asociarse con los AGs de grano grueso.

AGs Paralelos

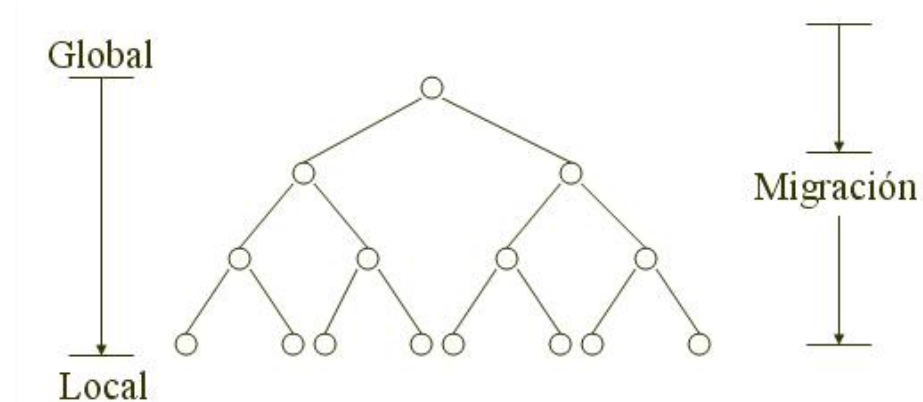
- En la arquitectura MIMD suele tenerse un número pequeño de **demes** (normalmente, menor a 40), pero el tamaño de cada uno de ellos suele ser grande.
- Es posible usar la misma representación para cada **deme**, o mezclar diferentes representaciones.

AGs Paralelos

- Los **demes** suelen diferenciarse debido al particionamiento del espacio de búsqueda.
- Las políticas de migración, en este caso, están dictadas por el propósito de los **demes**.

AGs Paralelos

Una topología que suele usarse con las arquitecturas MIMD es la de árbol, como se ilustra en esta figura:



AGs Paralelos

La migración en este caso introduce 2 nuevos parámetros:

- 1) ¿Con qué frecuencia exportar? (siempre se exporta el mejor). Si se hace con mucha frecuencia, se produce disrupción. Si se hace con poca frecuencia, habrá poca recombinación y puede producirse convergencia prematura en ciertos **demes**.

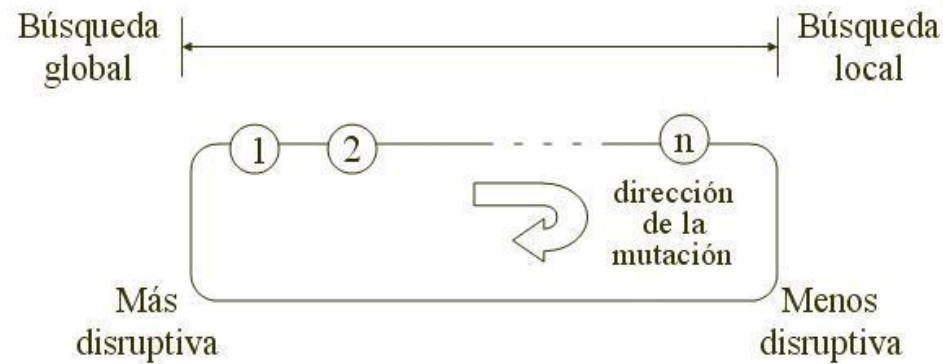
AGs Paralelos

2) ¿A qué **deme** exportar? Normalmente se usa una de las 2 siguientes opciones:

- Exportar el mejor individuo hacia el peor **deme**.
- Exportar hacia el **deme** donde se tenga una mejor correspondencia (*matching*) con respecto al individuo elitista, medida usando la distancia de *Hamming* (en el genotipo).

AGs Paralelos

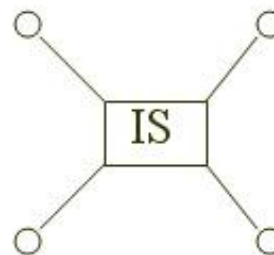
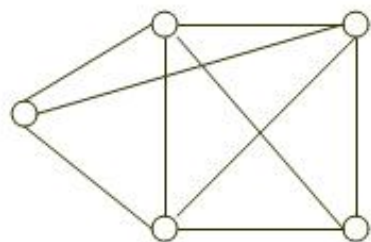
Otra posibilidad para una arquitectura MIMD es usar una topología de anillo:



En este tipo de topología, la búsqueda puede hacerse más local cambiando la precisión de la representación.

AGs Paralelos

Otra posibilidad es usar una topología de grafo con k interconexiones, o una topología de estrella:



En esta topología, se usan típicamente menos de 100 procesadores.

AGs Paralelos

Algunas opciones interesantes de la topología de grafo con k interconexiones (o de estrella) son las siguientes:

- 1) Mientras que la arquitectura es MIMD, pueden usarse la misma estructura y las mismas opciones para migración que con la arquitectura SIMD.
- 2) **Pizarrones**: Usando datos globales, cada **deme** decide por sí mismo cuándo cambiar su “dirección” de búsqueda.

AGs Paralelos

Otro punto interesante relacionado con los AGs paralelos son las **métricas**.

Normalmente, se consideran 2 de las métricas usadas con AGs simples:

- **Velocidad de convergencia:** Tiempo (generaciones) en alcanzar el óptimo.

AGs Paralelos

- **Precisión de la respuesta obtenida:** ¿Qué tan buena es la solución con respecto a la obtenida con otras técnicas?
- **Diversidad:** El grado en el cual los organismos (de una sola o de varias poblaciones) permanecen diferentes.

AGs Paralelos

Sin embargo, hay al menos una métrica adicional que es exclusiva de los AGs paralelos:

- **Velocidad de propagación de los esquemas:** ¿qué tan bien distribuidos están los esquemas “aptos”? O sea, ¿qué tan útil resulta la migración?

AGs Paralelos

Existen expresiones estándar para medir la diversidad de un AG (serial o paralelo). Consideremos por ejemplo la siguiente:

$$\delta = \frac{1}{l} \times \sum_{i=1}^l \left(1 - 2 \times \left| \frac{\sum_{j=1}^m \sum_{k=1}^n bit(i, k, j)}{m \times n} \right| \right)$$

l = Longitud cromosómica

m = número de **demes**

n = tamaño de cada **deme**

$bit(\cdot)$ = Valor del i -ésimo bit en el k -ésimo miembro del j -ésimo **deme**

δ = diversidad

AGs Paralelos

En esta fórmula, $\delta \in [0, 1]$ representa la diversidad de una población (o conjunto de poblaciones).

Si las cadenas consisten de puros ceros, $\delta = 0$.

Si las cadenas consisten de puros unos, $\delta = 0$.

Si las cadenas son del tipo 101010...10, $\delta = 0$.

AGs Paralelos

¿A qué se refiere la velocidad de propagación de esquemas?

- Se refiere no sólo al porcentaje de **demes** en los que un cierto esquema está presente, sino también al porcentaje en el cual dicho esquema está presente en un **deme** vecino.

AGs Paralelos

¿Cómo medimos la propagación de esquemas?

- Idealmente, deberíamos conocer de antemano cuáles son los buenos esquemas.
- Esto, sin embargo, es imposible en la práctica.

AGs Paralelos

Una alternativa viable es:

- Escoger varios esquemas de antemano.
- Hacer que la propagación de esquemas sea la fracción máxima de **demes** en la cual aparece un esquema.

AGs Paralelos

Ejemplo del cálculo de la propagación de esquemas (SP):

Esquemas seleccionados	% de demes en los que aparecen
*1*10*	3/9
*110**	4/9
*10*0*	5/9

En este caso: $SP = 5/9$