# INTRODUCCIÓN

## 1.1 ANTECEDENTES Y MOTIVACIÓN

El problema de corte en 2D restricto pertenece a la clase NP-Hard. Este es un problema muy interesante desde un punto de vista tanto teórico como práctico, que ha sido estudiado durante varios años, por lo que se ha generado mucho conocimiento relativo a él y muchas formas de abordarlo han sido propuestas. Sin embargo, debido a su fuerte dureza, métodos más eficientes para su resolución siguen siendo desarrollados en la actualidad, y mientras persista la interrogante P=NP, este seguirá siendo un problema abierto para el campo de la optimización.

Las capacidad de los AG celulares...

Existe evidencia de que un enfoque celular supera enfoques no estructurados en varios problemas académicos, como . Sin embargo no se han encontrado evaluaciones en problemas de carácter geométrico los cuales poseen una gran dureza, ampliamente reconocida ...donde las capacidades de ...Por otra parte existe evidencia de un enfoque distribuido supera el comportamiento de un AG secuencial, sin embargo comparaciones orientadas a ...

Sin embargo, hasta ahora

Por otra parte, en la literatura usualmente conocimiento específico del problema es utilizado para generar soluciones competitivas. Es interesante preguntarse si los conceptos celulares por si solos, permiten a la evolución genética mejorar el aprendizaje de reglas que permitan proveer resultados comparables a los obtenidos mediante dicho conocimiento específico. Lo cual podría ser observable en las características de las mejores soluciones evolucionadas.

Sin embargo para varias instancias simples ninguno de los AG fue capaz de alcanzar soluciones óptimas. Lo que lleva a pensar que un sesgo en la representación, incapaz de generar ciertas secuencias de piezas, impide que se visiten ciertas regiones del espacio, lo que al mismo tiempo no haya permitido al AG, y en especial al AG celular, alcanzar mejores soluciones.

Por otra parte, distintos strings binarios eran interpretados como las mismas secuencias de piezas por lo que esto también tiene efectos negativos en la diversidad de población, que usar directamente una representación de permutación para las secuencias de piezas.

## 1.2 DESCRIPCIÓN DEL PROBLEMA

Los problemas de corte y empaque pertenecen a una antigua y bien conocida familia, llamada *C&P* (cutting and packing). Esta es una familia de problemas naturales de optimización combinatoria, presentes en numerosas aplicaciones en el mundo real de la informática, ingeniería industrial, logística, fabricación, proceso de producción, etc. (Hifi and M’Halla, 2003). En términos abstractos, la estructura general de los problemas de *C&P* puede resumirse de la siguiente manera. Dados:

* Un conjunto de figuras pequeñas
* Un conjunto de regiones contenedoras

El objetivo es encontrar la mejor asignación posible (de acuerdo a algún criterio) de figuras pequeñas en las regiones contenedoras, respetando que:

* Las figuras se encuentren totalmente contenidas en las regiones
* Las figuras no se superpongan

Lo anterior implica varios problemas...akkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkkakakakjjaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

La complejidad de los problemas de C&P está fuertemente relacionada con la forma geométrica de los elementos que intervienen en la asignación. Así, la componente geométrica es el principal criterio para la clasificación de éstos (Beraudo et al., 2004).

A continuación se muestran algunos ejemplos de *C&P*. En la se muestran dos casos, en el primero hay una asignación de figuras regulares (rectángulos) y en el segundo una asignación de figuras irregulares (asimetrías y concavidades), ambas a una región contenedora rectangular. En la se muestra una asignación de figuras rectangulares a una región contenedora de largo infinito, (Ej.: corte de un rollo de género en industria textil), asignación a varias regiones contenedoras de distintas formas (Ej.: industria maderera, industria de vidrio) y asignación a regiones irregulares (Ej.: industria del cuero).



FIGURA 1.1 *Asignación de figuras regulares e irregulares a regiones rectangulares*



FIGURA 1.2 *Distintas* *formas de regiones contenedoras*

Existe una gran variedad de problemas de *C&P*, por lo que se han propuesto algunas tipologías, de manera de organizar y estructurar los problemas de una forma lógica, sin considerar sus aplicaciones ni disciplinas. (Dyckhoff, 1990) propone una tipología, la cual se resume en la , donde los problemas se dividen inicialmente en *2D* y *3D*, luego en regulares e irregulares, dentro de los regulares se encuentran las figuras rectangulares y otros tipos, y en los irregulares aparecen los polígonos y otros sin un patrón específico, finalmente los rectangulares se dividen en *no-guillotinables* y *guillotinables*.



FIGURA 1.3 *Clasificación de problemas de corte y empaque (Dyckhoff 1990)*

En particular, los problemas de *C&P* en *2D* son de gran relevancia en la producción y logística. Problemas de empaque en *2D* aparecen, por ejemplo, cuando varios artículos deben ser empacados en pallets en capas horizontales, o el posicionamiento eficiente de componentes en microchips. Problemas de corte en *2D* se encuentran en la customización de material en las industrias de vidrio, metal, madera y papel.

Dentro de los problemas de corte en *2D*, una variante clásica es el *TDC* (two-dimensional cutting) (Hifi and M’Halla, 2003), el cual considera una demanda de ítems rectangulares con tamaño y valor propios, que deben ser cortados de una placa rectangular, y tiene como objetivo determinar un patrón de corte (combinación de items a cortar), que maximice el valor total cortado. Si el valor de un ítem está dado por su área, el objetivo es maximizar el área total utilizada. Un patrón de corte, también llamado plan de corte, es considerado como factible si cada ítem se corta ortogonal (es decir, paralelo a los bordes de la placa), está completamente contenido en la placa, y si no hay superposición entre los ítems.

(Wascher et al., 2007) propone una tipología actualizada de problemas de *C&P*,la cual se ilustra en la . Según esta tipología, el *TDC* cae dentro del tipo de problema de *maximización de output* (Gonçalves and Resende, 2011), en donde el conjunto de regiones contenedoras no es suficiente para acomodar todos los items, por lo que todas ellas deben ser usadas. Por otra parte, los problemas de *minimización de input* son aquellos donde el conjunto de regiones contenedoras es suficiente para acomodar todos los items, luego todos los ítems deben ser producidos, utilizando la menor cantidad de regiones posibles. En este último tipo de problemas cae el denominado *CSP* (cutting stock problem), el cual describe el procedimiento de cortar el número requerido de items de un conjunto de placas rectangulares. El *TDC* puede ser usado como un problema auxiliar del *CSP*, ya que para resolver este último, el primero debe resolverse varias veces como un subproblema (Yoon et al., 2013).



FIGURA 1.4 *Clasificación de problemas de corte y empaque (Wascher 2007)*

Es importante señalar que, desatendiendo la operación de corte, la definición del *TDC* coincide con la del *2D-KP* (two-dimensional knapsack), en donde los ítems, en vez de ser cortados, deben ser empacados en un contenedor rectangular. De hecho, de acuerdo a (Beasley, 2000)los problemas de corte en *2D* también se conocen comúnmente como problemas de empaque, ya que se pueden ver como:

1. el problema de cortar piezas más pequeñas de una placa rectangular
2. el problema de empacar piezas más pequeñas en un contenedor rectangular

Debido a lo anterior, todas las consideraciones que se enuncien se pueden aplicar tanto al *TDC* como a su contraparte de empaque, *2D-KP.* El problema es fuertemente *NP-duro*, ya que en el caso especial en que todos los ítems tienen la misma altura, el problema de probar si todos ellos caben en el contenedor es equivalente al bien conocido *1D-BP* (one-dimensional bin packing) (Caprara and Monaci, 2003), el cual es fuertemente *NP-duro* (Garey and Johnson, 1979).

El tipo de ítem se define por sus dimensiones y por su valor. La demanda de ítems está dada por un conjunto de tipos de ítems. Con respecto al número de copias por tipo, de acuerdo a (Beasley, 2000) se distinguen las siguientes variantes:

* Irrestricto (Unconstrained) (*UTDC*): El número de copias por tipo no es fijo. Un patrón de corte puede por lo tanto, tener un número ilimitado de copias.
* Restricto (Constrained) (*CTDC*): Se fija un límite superior pi para el ítem i. Un patrón de corte puede por lo tanto, contener como máximo pi ítems de tipo i.
* Doblemente restricto (Doubly Constrained) (*DCTDC*): Se fija un límite superior pi y un límite inferior qj. Un patrón de corte puede contener como máximo pi y como mínimo qj items de tipo i.

Para complementar esta clasificación es interesante mencionar la clasificación de Fayard (Álvarez-Valdés et al., 2002), la cual además considera el valor de los items. Según esta clasificación se distinguen cuatro versiones para el *TDC*:

* Irrestricto no ponderado (Unconstrained unweighted) (*UU\_TDC*): El valor de cada ítem es equivalente a su área. Luego el objetivo es maximizar el área ocupada, equivalente a minimizar la pérdida.
* Irrestricto ponderado (Unconstrained weighted) (*UW\_TDC*): El valor de cada ítem es independiente de sus áreas. Luego el objetivo es maximizar el valor total de los ítems a cortar.
* Restricto no ponderado (Constrained unweighted) (*CU\_TDC*): Cada pieza tiene un límite superior de demanda bi, que restringe la cantidad de piezas de tipo i a cortar.
* Restricto ponderado (Constrained weighted) (*CW\_TDC*): Es el caso más general.

Desde un punto de vista práctico, el *UTDC* es un caso especial del C*TDC*, donde el número de veces que cada pieza puede aparecer en el patrón, está naturalmente restringido por el número de veces que cabe en la placa. Sin embargo el *UTDC* es generalmente más fácil de resolver que el C*TDC* (Cui and Huang, 2012). De hecho, a menudo es usado como un problema auxiliar del C*TDC* (Hifi and Zissimopoulos, 1997). En general, las variantes restrictas del *TDC* son más interesantes para las aplicaciones, y se ha dedicado mayor investigación a éstas (Álvarez-Valdés et al., 2005). Una explicación para esto es que, en la práctica, a menudo no se tienen cantidades ilimitadas de cada tipo de item, sino que estas están especificadas, por ejemplo en una orden de fabricación, o limitadas, por ejemplo por el inventario actual.

Por otra parte, si el valor de los items equivale a su área, la notación de Beasley es suficiente para describir el problema. Estos casos, correspondientes a las versiones no ponderadas, aparecen tipicamente en el corte de placas de acero o vidrio en piezas de ciertos tamaños requeridos, debiendo reducir las pérdidas al mínimo. Por su parte, las versiones ponderadas aparecen, por ejemplo, en una línea de producción, donde los valores podrían definir prioridades para ciertas piezas o incluso imponer que ciertas piezas, ya existentes en el patrón actual, deberían aparecer también en el siguiente patrón (Cung et al., 2000).

Con respecto al surtido de la demanda de ítems, es usual distinguir las variantes: homogénea (solo un tipo de ítem), débilmente heterogénea (pocos tipos de ítems y muchas copias por tipo), y fuertemente heterogénea (muchos tipos de ítems y pocas copias por tipo). De acuerdo a la tipología de Wascher, existe un *SLOPP* (Single Large Object Placement Problem) con una demanda débilmente heterogénea, y un *SKP* (Single Knapsack Problem) con una demanda fuertemente heterogénea. La figura 1.5 muestra esta clasificación.



FIGURA 1.5 *Landscape de tipos de problemas intermedios: maximización de output (Wascher 2007)*

Del total de publicaciones revisadas, pocas identifican el *TDC* dentro de la tipología de Wascher. En particular, (Chen, 2007) señala que el *TDC* es conocido formalmente como *SLOPP*, no obstante, (Bortfeldt and Winter, 2008) señalan que para las variantes *CTDC* y *DCTDC* pueden haber ambos *SLOPP* y SKP, en tanto que para la variante *UTDC* solo debería ser asumido *SLOPP*. Por lo tanto una revisión de publicaciones en torno al *TDC* debería considerar tanto *SLOPP* como *SKP*. Para una revisión de publicaciones y otros recursos relacionados a los problemas de *C&P* el lector puede dirigirse al sitio de [*ESICUP*](http://paginas.fe.up.pt/~esicup/tiki-index.php).

Adicionalmente se añaden las siguientes restricciones, las cuales provienen de las limitaciones prácticas y/o tecnológicas de la aplicación en cuestión:

* Restricción de orientación: Fija la orientación de todos los ítems, y prohíbe su rotación. Esto implica que un ítem de dimensiones (a,b) es distinto a un ítem de dimensiones (b,a).
* Restricción de corte de guillotina: Exige que los cortes sean realizados de lado a lado sin parar, ya sea horizontal o verticalmente, tal como se puede ver en la figura 1.6.



FIGURA 1.6 *Ejemplos de patrón guillotinable y no-guillotinable*

La restricción de orientación aparece, por ejemplo, por la calidad de la superficie del material (como resultado de la laminación), también en el diseño de páginas de periódicos, donde los elementos deben tener una orientación fija. Cabe señalar que el tratamiento de ítems con rotación es rara vez encontrado y una restricción de orientación es casi siempre asumida. Sin embargo (Bortfeldt and Winter, 2008) señalan que desde el punto de vista sistemático el caso sin una restricción de orientación (con rotación) es primario, mientras que el caso con una restricción de orientación (sin rotación) representa un problema derivado y por lo tanto secundario. Por esta razón, las variantes de problemas sin esta restricción deberían ser considerados.

Por otra parte, la restricción de guillotina aparece, por ejemplo, cuando se utilizan hojas de guillotina para dividir la placa en piezas rectangulares, lo cual es muy común en las industrias de fabricación. Igualmente, en una situación de empaque, a menudo se requiere que los elementos se puedan descargar en etapas, extrayendo simultáneamente todos los artículos empacados en la misma vertical u horizontal.

En la práctica, maximizar el valor cortado o empacado a menudo no es suficiente. También son importantes otros aspectos, como los tiempos de operación, o la vida útil de la maquinaria utilizada. (Cui and Huang, 2012) señalan que la elección del patrón debe considerar el compromiso entre dos aspectos: utilización del material y complejidad del patrón; generalmente patrones simples conducen a una baja utilización del material, por el contrario patrones más complejos conducen a una mejor utilización del material.

Debido a lo anterior, formulaciones más orientadas a las aplicaciones, imponen una restricción adicional, que consiste en restringir el número de etapas a una cantidad acotada, de manera de simplificar el patrón y así reducir los tiempos de operación. Esta restricción se conoce como k-etapas, donde k denota el número máximo de etapas permitido para realizar la operación de corte o descarga. Típicamente, esta restricción fija el número de etapas a 2 o 3, con lo que se obtienen patrones simples aunque en ocasiones con baja utilización del material. La figura 1.7 muestra que 4 etapas son necesarias para obtener el patrón de la figura 1.6. En la Figura 1.8 se muestra un ejemplo de patrón en 2-etapas y un patrón en 3-etapas.



FIGURA 1.7 *Número de etapas necesarias para obtener el patrón de la*



Figura 1.8 *Ejemplos de patrones en 2-etapas y 3-etapas*

Desde un punto de vista computacional, la restricción de k-etapas simplifica el problema (Dolatabadi et al., 2012). Por ejemplo, es posible formularlo como un problema de programación lineal extendiendo la formulación del bien conocido 1D-CP. Propuesta por ...Hasta la fecha no se conocen formulaciones matemáticas para la variante sin etapas, por lo que métodos de enumeración o métodos de búsqueda. Es por ello que, en términos de dureza, resulta más interesante el problema con patrones de guillotina generales.

La variante que se va a abordar en este trabajo de titulación es el *CTDC* con las restricciones de orientación y guillotina. Este es uno de los problemas más interesantes, ya que es computacionalmente más difícil de resolver que cualquier otra variante de *TDC* (por ejemplo es más difícil que el *UTDC* o el *TDC* en k-etapas). (K. Yoon 2012) El problema puede ser caracterizado de la siguiente forma:

Las dimensiones de la placa *R* son *L* *x* *W* (length y width); el *i*-ésimo item tiene dimensiones *li* *x* *wi*, valor vi (equivalente a su área *si=li\*wi*), y un límite superior de demanda *bi*, *i=1,...,m*. El objetivo es cortar la placa en *xi* piezas del tipo *i*, de manera que *0<=xi<=bi, i=1,...,m* y la utilidad total sea maximizada. Además se asume que:

1. Todos los items tienen orientación fija
2. Todos los cortes aplicados son de tipo guillotina
3. Los parámetros *L*; *W*, *li*, *wi*, y *bi*, i=1,...,m, son enteros no negativos

El vector (x1,...,xn) corresponde a un patrón de corte, si es posible producir xi piezas de tipo i, i=1,...,m, en la placa R sin superposición.

Tanto la fuerte dureza inherente del *CTDC* como sus variadas aplicaciones, lo convierten en un problema muy interesante, por lo que la investigación de nuevas técnicas y métodos más eficientes para abordarlo continúan siendo desarrollados en la actualidad. Así mismo, constituye un buen candidato

## 1.3 SOLUCIÓN PROPUESTA

### 1.3.1 Características de la solución

La solución propuesta consiste en modelar el problema de corte de piezas mediante un enfoque evolutivo. Existen varias formas para hacer esto, sin embargo, dado que se desea evaluar el comportamiento en la búsqueda sobre el espacio de soluciones, el modelo debe estar pensado para permitir que el algoritmo pueda navegar de forma libre, guiándose solamente por la evolución, tanto en un contexto de no estructuración como estructuración de población. Para promover la reducción del sesgo en la búsqueda de soluciones, se adopta la representación propuesta por (Flores, 2012), la cual está basada en un esquema genotipo-fenotipo, en donde se abstrae la interpretación del cromosoma del nucleo de procesamiento genético. Esta representación está pensada para que el algoritmo...

Para determinar si la estructuración de la población mediante un enfoque celular permite a un AG obtener resultados más competitivos versus el uso de otros enfoques conocidos, se han seleccionado distintas familias de AG canónicos: dos familias con poblaciones no estructuradas: AG generacional y AG steady-state y una familia de AG descentralizado conocida como AG distribuido, en el cual la población es particionada en varias subpoblaciones separadas que cooperan entre sí. Luego se debe diseñar cada uno de los AG propuestos, seleccionando los operadores genéticos más apropiados para el problema en cuestión dada la representación adoptada.

### 1.3.2 Propósito de la solución

El propósito de este trabajo es evaluar el desempeño computacional de un algoritmo genético cuya población se estructura bajo conceptos celulares, y de esa manera, determinar si este enfoque es más apropiado que un enfoque no estructurado, para tipos de problemas de carácter geométrico

## 1.4 OBJETIVOS Y ALCANCES DEL PROYECTO

### 1.4.1 Objetivo general

Estudiar los efectos de la estructuración de la población en un AG mediante un enfoque celular, aplicado a problemas de carácter geométrico, en particular para el problema de corte de piezas guillotinable bidimensional restricto

### 1.4.2 Objetivos específicos

1. Realizar un estado del arte sobre algoritmos genéticos celulares
2. Realizar un estado del arte sobre algoritmos genéticos aplicados al problema en estudio o similares
3. Diseñar e implementar las representaciones y funciones constructoras para el problema en estudio
4. Diseñar e implementar un AG, un AG distribuido y un AG celular que resuelvan el problema en estudio
5. Identificar conjuntos de instancias de prueba
6. Diseñar un experimento computacional que permita evaluar los algoritmos propuestos
7. Analizar los resultados obtenidos

### 1.4.3 Alcances

En este trabajo diferentes modelos de poblaciòn en AG son comparados en términos de su aproximaciòn a las soluciones óptimas para distintas instancias del problema en estudio, para ello, se diseña cada uno de los AG a evaluar, de manera apropiada en funciòn del a. Además se verifican los efectos que tiene estructurar la población, sobre la bùsqueda que realiza el AG.

## 1.5 METODOLOGÍAS Y HERRAMIENTAS UTILIZADAS

### 1.5.1 Metodología y herramientas a utilizar

En este trabajo se utiliza una plataforma computacional evolutiva, JCELL , para evolucionar soluciones para el CTDC. Se realiza la evoluciòn bajo cuatro enfoques: 2 no estructurados, concistentes en AG generacional y AG steady-state y 2 estructurados, concistentes en AG distribuido y AG celular. Se utilizan 49 instancias ampliamente usadas en la literatura, cuyos òptimos son conocidos. Así, se evalúa la calidad de las soluciones alcanzadas por los distintos enfoques, mediante la precisiòn o su proximidad a la solución óptima. Para el diseño y configuraciòn de cada algoritmo, se usa una plataforma de optimización de algoritmos, ParamILS, lo cual es complementado por el anàlisis del problema en conjunto con la representación seleccionada. Para el experimento se realizan 30 ejecuciones por cada algoritmo para cada instancia, lo cual nos permite asumir una distribuciòn normal de los datos, y considerar el promedio de las 30 ejecuciones como el valor representativo. Para sustentar las hipòtesis de trabajo se realizan pruebas estadísticas t-student, cuando se realizan comparaciones entre 2 algoritmos y ANOVA de un factor cuando se realizan comparaciones entre màs de 2 algoritmos. Los test estadìsticos son realizados mediante la suite estadìstica IBM SPSS.

### 1.5.3 Ambiente de desarrollo

## 1.6 ORGANIZACIÓN DEL DOCUMENTO

El Capítulo 2 presenta la revisión de la literatura respecto al problema de corte de piezas bidimensional restricto en un contexto general y desde un enfoque evolutivo. En él, quedan al descubierto las respectivas brechas del conocimiento que justifican la realización de esta Tesis. El Capítulo 3, describe el procedimiento experimental realizado. En éste se enuncian los conceptos de estructuración bajo un enfoque celular y como éstos fueron aplicados a la generación automática de algoritmos. En el Capítulo 4, se detallan los resultados obtenidos para determinar un tamaño mínimo de población y se analiza la eficiencia y calidad del proceso con respecto a las principales variables de interés. El Capítulo 5, presenta las respuestas a las dos preguntas de investigación, mediante el análisis de los resultados obtenidos al usar distintos conjuntos de entrenamiento o casos de fitness y la robustez de los algoritmos frente a instancias grandes. Finalmente, el Capítulo 6 presenta las conclusiones del estudio, donde se resumen los principales hallazgos y se enumeran las consideraciones para trabajos futuros.

## 1.7 GLOSARIO

# ESTADO DEL ARTE

## 2.1. EL PROBLEMA DE CORTE DE PIEZAS GUILLOTINABLE BIDIMENSIONAL RESTRICTO (CTDC)

El estado del arte del *TDC* es muy extenso y variado; en la literatura existen diversos enfoques, los cuales se pueden clasificar en: exactos, heurísticos y metaheurísticos. Los enfoques exactos están representados principalmente por *tree-search* o *branch-and-bound*; también existen métodos exactos basados en *dynamic optimization* y otros son basados en modelos. Todos ellos también pueden ser encontrados en métodos heurísticos, sin garantía de optimalidad. Los enfoques metaheurísticos incluyen algoritmos genéticos (*AG*), Simulated Annealing (*SA*), Tabú Search (*TS*) y Greedy Randomized Adaptive Search Procedure (*GRASP*).

Dentro de los métodos exactos varios estudios se han enfocado en clases especiales de patrones de corte, tales como los: *p-group*, *k-staged* (Scheithauer 2004, Alves 2010), *t-shaped*. Sin embargo, pocos estudios han considerado algoritmos exactos para el *CTDC* (sin etapas). Existen dos enfoques algorítmicos para abordar este problema: *top-down* y *bottom-up*. El enfoque *top-down* genera todos los patrones posibles mediante el corte sucesivo de subplacas. Así, el problema se define como la búsqueda en un árbol como el que se muestra en la figura 2.1, donde las ramas representan cortes horizontales o verticales, y los nodos representan la subplaca resultante del corte.

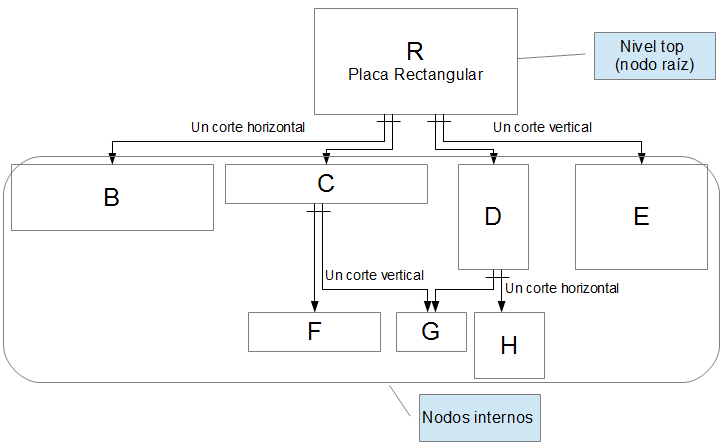


FIGURA 2.1 *Enfoque top-down: Proceso de búsqueda en el grafo dirigido*

En contraste, el enfoque *bottom-up*, se basa en la observación de que cualquier patrón que satisfaga la restricción de guillotina, puede ser obtenido mediante la construcción horizontal o vertical de rectángulos, como se muestra en la figura 2.2. Todas las posibles combinaciones de rectángulos más pequeños son generadas para obtener rectángulos más grandes hasta que no puedan obtenerse más patrones de guillotina. Ambos enfoques pueden usar *upper* y *lower-bounds* en procedimientos de búsqueda, para descartar ramas no prometedoras.

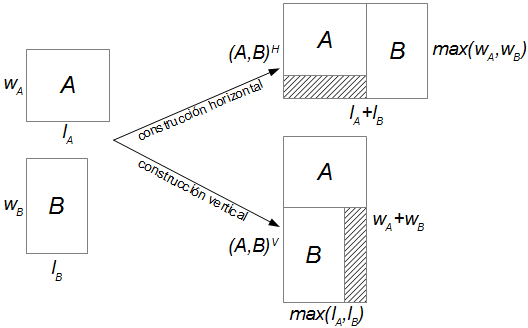


FIGURA 2.2 *Enfoque bottom-up: Construcción horizontal o vertical de patrones*

El *CTDC* ha sido resuelto de manera exacta mediante el uso de ambos enfoques. (Christofides and Whitlock, 1977) propusieron originalmente el enfoque *top-down*; desarrollaron un procedimiento *tree-search* basado en una estrategia de búsqueda *depth-first*, resolviendo instancias pequeñas de manera exacta. Además proponen un procedimiento de discretización o normalización, el cual considera solamente combinaciones lineales de las dimensiones de las piezas ordenadas. Esto permite lidiar con un número finito de cortes en métodos de enumeración. (Hifi and Zissimopoulos, 1997) propusieron un algoritmo mejorado, el cual utilizaba *lower* y *upper-bounds* más efectivos. Aplicaron una estrategia de búsqueda *depth-first*, la cual requiere una pequeña cantidad de memoria, a expensas de mayor tiempo computacional y dificultades en presencia de restricciones adicionales. El enfoque *bottom-up*, propuesto originalmente por (Wang, 1983), fue generalizado por (Viswanathan and Bagchi, 1993); desarrollaron un método *branch-and-bound* basado en una estrategia de búsqueda *best-first*, donde cada nodo corresponde a un patrón de corte *R* cuya área es *g(R)* y *P* el área en la placa sin ocupar, (ver figura 2.3). La estrategia *best-first* consiste en expandir el nodo más prometedor, utilizando para ello, una estimación del *upper-bound* relativo a *P[[1]](#footnote-1)*. Se generan combinaciones verticales u horizontales del patrón seleccionado con los patrones correspondientes a los nodos hoja. (Hifi, 1997) mejoró este algoritmo mediante la aplicación de *lower* y *upper-bounds* más efectivos. (Cung et al., 2000) mejoró su eficiencia añadiendo algunas estrategias de poda y almacenando "códigos de patrones" en los nodos, para remover patrones repetidos, los cuales comparten el mismo tamaño y la misma combinación de piezas. *best-fist* requiere más memoria, pero menor tiempo de cómputo, además es flexible ante restricciones adicionales. Recientemente (Yoon et al., 2013) proponen varias mejoras al algoritmo anterior; un nuevo método más eficiente es usado para remover patrones duplicados, se usa una estrategia eficiente de poda, se proponen dos nuevos *upper-bounds* y se proponen dos métodos para prevenir la formación de patrones dominados. El algoritmo es comparado con el de (Cung et al., 2000), el cual había sido previamente el algoritmo exacto más eficiente para el *CTDC*. Además para las instancias de escalas mayores se incluye una comparación con el algoritmo heurístico *TDH*, el cual está basado en un enfoque *top-down* con *hill-climbing*. El algoritmo propuesto reduce de manera dramática la cantidad de nodos generados en el árbol de ramificación, gracias a técnicas eficientes de poda y a *upper-bounds* más estrictos. Además obtiene soluciones óptimas para instancias de larga escala, previamente desconocidas.



FIGURA 2.3 *Branch-and-bound: estado de un nodo; tamaño de del patrón R y la placa*

Dentro de los métodos basados en modelos, un enfoque novedoso es el que propone la teoría de grafos. Recientemente (Clautiaux et al., 2013) presentan un modelo teórico llamado *grafo de guillotina*. Este modelo usa grafos de coloración dirigidos, en donde los circuitos están relacionados a combinaciones horizontales o verticales. La idea es que cada *grafo de guillotina* puede ser asociado con una clase específica de patrón, llamada *clase corte-guillotina*, esto permite enfocarse en subconjuntos dominantes de soluciones, evitando redundancias en métodos de búsqueda.



FIGURA 2.4 *Algunos elementos de una Clase Corte-Guillotina*

Las *clases corte-guillotina* describen conjuntos de patrones equivalentes (ver figura 2.4), mediante una expresión denominada Recursive Multi Build (*RMB*), la cual se define como: un solo ítem, o una composición de combinaciones verticales horizontales sucesivas de *RMB*. Dicha expresión permite representar una gran cantidad de secuencias de combinaciones, cuyos patrones resultantes comparten la misma estructura combinatoria y pueden ser considerados como la misma solución. Una implementación computacional eficiente para representar los *RMB* es mediante grafos bi-coloreados dirigidos, en donde cada nodo corresponde a un *RMB* y cada arco, según su color, representa una combinación vertical u horizontal entre *RMB*. En este grafo los ciclos monocromáticos representan *RMB* que pueden ser reducidos mediante una contracción del circuito. Inicialmente cada nodo corresponde a un *RMB* compuesto de un solo ítem, los cuales se van reduciendo en sucesivas contracciones de circuitos monocromáticos. Si es posible reducir el grafo a un único nodo se habla de un grafo de guillotina (ver figura 2.5).



FIGURA 2.5 *Modelando la Clase Corte-Guillotina asociada al patrón con un grafo de guillotina*

El modelo es embebido en un esquema de *Constraint Programming*. El cual busca un conjunto adecuado de arcos. *Constraint Programming* es un paradigma orientado a problemas combinatoriales que pueden ser descritos por un conjunto de variables, un conjunto de valores para cada variable, y un conjunto de restricciones entre las variables. Su potencial radica en un proceso denominado “propagación de restricción”, donde se realizan ciertas deducciones que reducen el esfuerzo computacional. El algoritmo es evaluado en el problema de strip cutting, en el cual el ancho es fijo..................

En general los algoritmos exactos no son un método práctico para resolver el *CTDC*, debido a la complejidad y a la explosión combinatoria, luego el costo computacional se torna exorbitante para instancias de mayor dureza. En estos escenarios, a menudo se usan heurísticas. Dentro de los métodos heurísticos, (Wang, 1983) Propuso un algoritmo constructivo usando el ya mencionado enfoque *bottom-up,* el cual genera patrones agregando sucesivamente piezas o grupos de piezas (soluciones parciales), generando así, nuevas soluciones parciales. Para evitar la explosión combinatoria, se eliminan las soluciones duplicadas y se define un criterio de "aspiración" para cada solución parcial, mediante la fijación de un parámetro *β,* el cual debe satisfacer *Tp<βWH*, donde *Tp* es la pérdida total de un patrón *P*. Este parámetro permite establecer un balance entre tiempo de computación y garantía de optimalidad. Posteriormente el algoritmo fue mejorado independientemente por (Vasko, 1989), descartando soluciones parciales que no puedan seguir siendo combinadas vertical u horizontalmente, y por (Oliveira and Ferreira, 1990) el cual cambia el criterio de "aspiración" para rechazar lo antes posible las soluciones parciales con pérdida total superior a *βWH*. Para ello, se considera un *lower-bound* asociado a la pérdida externa de la solución parcial en conjunto con la pérdida interna, para decidir sobre la aceptación de la solución parcial. (Morabito and Arenales, 1992) presentan un enfoque basado en el modelo *And/ord-Graph* (grafo y/o), el cual representa cada posible patrón como un camino completo en un grafo y/o, donde los nodos corresponden al rectángulo inicial, los rectángulos intermedios, las piezas, o las pérdidas, y los arcos corresponden a los cortes a los rectángulos. Propusieron una búsqueda *depth-first*, usando solo cortes normalizados y aplicando reglas de simetría, orden en los cortesy estrategias *hill-climbing* para reducir la búsqueda. En este trabajo resuelven instancias de larga escala para el *UTDC*. Unos años más tarde, (Morabito and Arenales, 1996) extendieron su enfoque previo, al *CTDC*,obteniendo buenos resultados. (Fayard et al., 1998) diseñaron un algoritmo basado en la resolución de una serie de problemas de la mochila usando *dynamic programming* para el *UTDC.* Además mostraron cómo su enfoque también podía ser usado para resolver de manera aproximada el *CTDC*. (Álvarez-Valdés et al., 2002) desarrollaron varios algoritmos heurísticos de propósito general para resolver las cuatro variantes del *TDC.* Propusieron dos procedimientos constructivos basados en limites simples mediante la resolución del problema de la mochila unidimensional. Usaron dichos algoritmos constructivos como *building-blocks* para procedimientos más complejos. Desarrollaron un algoritmo *GRASP* y un algoritmo *TS*. De este trabajo destaca *TS,* el cual es capaz de obtener resultados de alta calidad en tiempos moderados. (Hifi, 2004) presentó un algoritmo híbrido (denominado *TDH*) para el *CTDC*, en el cual, una búsqueda *depth-first* usando estrategias *hill-climbing* y *dynamic programming* son combinados. El algoritmo puede producir buenas soluciones para las instancias de larga escala. (Cui, 2007) presentó dos algoritmos exactos para el *CTDC.* basados en *branch-and-bround* combinado con técnicas de *dynamic* programming: uno para generar patrones *T-shape* para simplificar el proceso de corte, el otro para generar patrones en *3-etapas homogéneos*. Muestran cómo estos algoritmos pueden ser usados como heurísticas para generar patrones generales para el *CTDC*. (Chen, 2007) presentó un algoritmo heurístico recursivo (*REC*) para el *CTDC*, cuya formulación se basa en ubicar las piezas en la esquina inferior izquierda del bloque, de lo cual se obtienen dos formas de dividir la región inutilizada, vertical u horizontalmente. Luego el valor del bloque equivale al valor de la pieza ubicada más el valor de las pérdidas producidas (que también son bloques). Para reducir los tiempos computacionales usa *upper y lower-bounds*, se ordenan las piezas de acuerdo a un orden decreciente de sus valores y se restringe el tiempo computacional destinado al bloque de prueba actual. El algoritmo *REC* produce buenas soluciones en corto tiempo para problemas de varias escalas. (Morabito and Pureza, 2008) proponen un método heurístico para el *CTDC*, el cual usa una relajación del espacio de estados de una formulación mediante dynamic programming para el *CTDC* en *k* etapas. Proponen un algoritmo de optimización subgradiente, el cual fija un valor de *k* suficientemente grande, y el cual incluye una heurística interna que convierte soluciones no factibles dadas en un paso dado del algoritmo, en soluciones factibles. Utilizan un enfoque And/or-graph para inicializar el *lower-bound* al comienzo del algoritmo. Se concluye que el algoritmo es competitivo comparado a otros métodos propuestos en la literatura y requiere cortos tiempos computacionales para proveer la mejor solución. Recientemente, (Cui and Chen, 2012) proponen una heurística simple, basada en una clase de patrón denominada *patrón* *de bloque extendido*, el cual contiene un número *k* de piezas del mismo tipo (piezas principales), ubicadas en la esquina inferior izquierda de la subplaca correspondiente. Se establecen 4 casos de *patrón de bloque extendido* (ver ), donde las piezas principales forman una fila o una columna y si el área no ocupada en la placa se divide en forma horizontal o vertical, respectivamente. Se propone un algoritmo goloso, denominado *HCEB* (heuristic for constrained extended blocks patterns), el cual va construyendo la solución de acuerdo a la fórmula de recursión *F(x,y)*, equivalente al maximo entre *F(x-1,y)*, *F(x,y-1)* y los valores correspondientes a cada unos de los *patrones de bloque extendidos*, previamente explicados. El valor inicial de *F(x,y)* es el mayor entre *F(x-1,y)* y *F(x,y-1).* y los valores de los 4 tipos de patrones antes descritos, los cuales son considerados para mejorar la solución. Antes de considerar un tipo de patrón se verifica si su valor es superado por un *upper bound* relativo y si la frecuencia de la pieza i no excede la restricción del problema. Si estas condiciones se cumplen se intenta mejorar el patrón mediante una reasignación de piezas a subplacas. *HCEB* opera de la siguiente manera, primero inicializa F(x,y) y n(x,y,j) a 0 para x e y inferiores a la dimensiones mínimas de las piezas. Las dimensiones de las subplacas se van enumerando de acuerdo a un orden ascendente del tamaño de las piezas. Luego para cada subplaca el algoritmo evalúa F(x,y) para cada pieza disponible.



FIGURA 2.6 *Patrón de bloque extendido: los cuatro casos*

*HCEB* es comparado con otros tres algoritmos basados en enfoques heurísticos: *TS500* un algoritmo basado en búsqueda tabú, *TDH2* un algoritmo que usa programación dinámica y hill climbing, y *REC* un algoritmo recursivo. Para instancias de larga escala, *HCEB* es superior al resto. Se concluye que el algoritmo provee un buen balance entre calidad de solución y tiempo computacional. Por otra parte, debido a su simpleza es sencillo de implementar. Además puede ser usado como una buena solución inicial para algoritmos exactos.

## 2.2. ALGORITMOS GENËTICOS APLICADOS AL PROBLEMA DE CORTE DE PIEZAS GUILLOTINABLE BIDIMENSIONAL RESTRICTO

Debido a la relevancia que tienen para este estudio, se dedica esta sección para exclusivamente presentar los principales avances en este tipo de problemas desde el enfoque evolutivo. Estableciendo así, un referente sobre cómo el problema ha sido abordado mediante esta metodología, así como los principales resultados, los cuales serán de vital importancia, tanto para el diseño del AG, como para el análisis de los resultados obtenidos.

Según (Hopper, 2000) existen básicamente tres enfoques para abordar el problema. El más común es el enfoque en dos etapas, donde el AG es usado para explorar y manipular el espacio de soluciones, dadas por secuencias de elementos, y un segundo procedimiento, consistente en una rutina de colocación, es usado para evaluar las soluciones generadas. Por ejemplo (Leung et al., 2003) aplicaron un algoritmo genético simple (*GA*) al 2D-KP. Usaron un enfoque en dos etapas, en donde una representación de permutación, correspondiente a la secuencia de piezas que debe ser empacada/cortada, constituye el genotipo, y una estrategia de colocación (*DP*) es usada como algoritmo decodificador, para así obtener un *layout* válido, o fenotipo. Este trabajo, también constituye una evidencia de que se ha intentado mejorar el comportamiento en la búsqueda del algoritmo para este problema, ya que la principal motivación fue la de intentar aliviar el problema de *convergencia prematura*. En su investigación, ellos encontraron que *GA* usualmente producía buenos resultados, sin embargo convergía muy rápido, de manera que buenos resultados eran producidos en etapas tempranas de la evolución. Propusieron un algoritmo genético mixto con simmulated annealing (*SAGA*), el cual fue comparado con *GA*. Introdujeron un operador que induce competencia entre los padres y los hijos; si los hijos son mejores que sus padres, estos se aceptan, pero si no, aún pueden aceptarse con cierta probabilidad. Hicieron variar esta probabilidad según la temperatura en un enfoque *simulated annealing*. Observan que *SAGA* se comporta un poco mejor que *GA*, porque, mientras que *GA* converge a buenas soluciones, este rápidamente se homogeniza, en tanto que *SAGA* inicialmente mantiene peores soluciones, sin embargo estas son capaces de evolucionar a individuos superiores en un momento maduro de la evolución (ver figura 2.7).



FIGURA 2.8 *GA vs SAGA Pérdida promedio en la población versus Iteraciones*

Un problema aparente del enfoque en dos etapas es su fuerte dependencia del decodificador, ya que el conocimiento del dominio está oculto en la rutina de colocación. Por otra parte, un decodificador podría limitar al AG, al no soportar la herencia de ciertas características en la descendencia. (Hopper and Turton, 2000) realizan una investigación empírica de algoritmos heurísticos y metaheurísticos para el *2D-KP*. Consideran dos heurísticas: bottom-left (*BL*), la cual es simple y eficiente en tiempo, y bottom-left fill (*BLF*), la cual es capaz de rellenar espacios pero a costa de mayor tiempo computacional. Estas heurísticas son hibridizadas con tres metaheurísticas: algortimo genético *AG*, simulated annealing *SA*, naïve evolution *NE* y una heurística de búsqueda local (*hill-climbing*). Su estudio compara los algoritmos híbridos en términos de calidad de solución y tiempo computacional. Además, con el fin de mostrar la efectividad de estos, su rendimiento es comparado con búsqueda aleatoria y rutinas heurísticas de empaque. Sus resultados muestran que en términos de calidad de solución, las metaheuristicas superan a las heurísticas, siendo *SA* la mejor de todas. Sin embargo *GA* y *NE* son mejores en términos de tiempo computacional. Debido a que las diferencias en el desempeño entre los algoritmos híbridos usando BL y BLF son debidas a la heurística mejorada, el decodificador tiene mayor efecto en el desempeño de la técnica híbrida que la metaheurística en sí. Esto parece sugerir enfoques donde mayor conocimiento del layout sea incorporado en la metaheuristica en vez del decodificador. Sin embargo señalan que este último tipo de representación, estudiado por otros investigadores no necesariamente logró mayores densidades empacadas que las técnicas híbridas para los problemas probados.

Un segundo enfoque intenta incorporar más información acerca del layout dentro de la estructura de datos del AG, aunque se necesitan ciertas reglas adicionales para fijar la posición en el layout. Por ejemplo, (Ono and Ikeda, 1998) proponen un GA que utiliza una parte binaria, representando los operadores binarios V o H, correspondientes a combinaciones verticales u horizontales respectivamente, y una parte numérica, representando las piezas. Así, el cromosoma es interpretado como un árbol binario, leído en notación post-fija (ver figura 2.7). Se imponen ciertas restricciones sobre la cantidad de genes binarios respecto a genes numéricos y su posición relativa en el cromosoma para representar una solución válida para el problema. De acuerdo a la ecuación, todas las piezas son integradas recursivamente mediante un algoritmo de layout, utilizando el enfoque *bottom-up*



FIGURA 2.7 *Cromosoma y su interpretación como layout guillotinable*

Otro caso es (Beasley, 2000) quien utiliza una representación binaria/real. La parte binaria indica si la p-ésima copia de la pieza es cortada o no de la placa. Mientras que la parte real indica las coordenadas del centro de la pieza correspondiente, las cuales son aproximadas al entero más cercano. El *fitness* está dado por dos medidas: *fitness* es la función objetivo original, y *unfitness* indica en qué grado, un individuo está violando la restricción de superposición. Unos años después (Beraudo et al., 2004) adoptan la representación de (Beasley, 2000), con la diferencia de que las coordenadas no son aproximadas a enteros. El *fitness* está dado por la diferencia entre las piezas a cortar que pudieron ser colocadas en el layout y las piezas a cortar que no pudieron ser colocadas, y es usado como un mecanismo de penalización. El objetivo es maximizar dicha diferencia, ya que una solución que contenga piezas que no pudieron ser cortadas, es menos deseable que otras donde todas las piezas hayan podido ser cortadas.

A diferencia de los dos primeros enfoques, el tercer enfoque traslada el proceso de búsqueda genética al dominio de layouts. Debido a que las operaciones genéticas son realizadas directamente en los layouts, este método no requiere una técnica de decodificación. Tiene la ventaja de que la implementación de los conceptos metaheurísticos (vecindad, operadores genéticos) tienen una significación y aplicación directa para el problema. Sin embargo, a diferencia de un enfoque en dos etapas, donde las restricciones geométricas son aplicadas por un algoritmo decodificador, en este enfoque la representación debe hacerse cargo de las restricciones. Usualmente alguno de los siguiente mecanismos son usados para ello: rechazo, reparación o penalización. Un ejemplo de este enfoque es (Bortfeldt and Winter, 2008), quienes introducen un algoritmo genético llamado Strip Packing Genetic Algorithm Layer (*SPGAL*). El *AG* propuesto no usa una codificación especial para las soluciones, en lugar de ello, la búsqueda tiene lugar directamente en el espacio fenotípico, dado por planes de empaque. El *AG* genera planes con una estructura de capas, donde el largo de cada capa está dado por el largo de una pieza definitoria, de manera que cada pieza se encuentra completamente contenida en una capa (ver figura 2.8). Para medir la calidad de una capa, utilizan el filling rate (*fr*), que equivale al cuociente entre la suma de todas las áreas de las piezas contenidas en una capa y el área de la capa. Mientras que el *fitness* equivale al área total de todas las piezas empacadas.



FIGURA 2.9 *Estructura de capas de un plan de empaque*

Para generar la población inicial utilizan la heurística Best Fit Decreasing Height (*BFDH*) con algunas mejoras. Para el cruzamiento, las capas con mejores *fr* son extraídas de los padres para combinar las buenas características de estos. Utilizan un procedimiento heurístico de completado de capas, el cual consiste en un algoritmo *tree search* que agrega capas y en cada paso selecciona la que tiene mayor *fr*. Además consideran una post-optimización para mejorar aún más la mejor solución encontrada al final de la búsqueda genética. Esta reduce las pérdidas resultantes en los bordes de las capas. Aplican un enfoque de diversificación, en donde la búsqueda genética es repetida varias veces bajo distintas condiciones y la mejor solución obtenida es emitida al final. El *AG* propuesto fue comparado con varias heurísticas recientes de esa época, sus resultados mostraron que éste, generalmente mostraba un comportamiento medio para el caso con restricción de guillotina, lo cual atribuyeron a la estructura adicional de capas de los planes de empaque. Sin embargo, para instancias de larga escala sin la restricción de guillotina el AG propuesto alcanza la calidad de las mejores heurísticas conocidas.

(Wang, 2010) propone un AG adaptivo para el *2D-KP*. Su enfoque es distinto al resto, ya que formula el problema como un problema de *factores atractivos dinámicos*, donde, dadas *n* piezas que están esperando en el layout, el objetivo es poner tantos bloques rectangulares como sea posible. El requisito es que la tasa de utilización del área sea la mayor posible y que los bloques rectangulares no se superpongan. Se define una función de posicionamiento de bloques, cuyos diversos parámetros, determinan diferentes ubicaciones de los bloques rectangulares, lo cual genera diferentes layouts. Así, el problema es transformado a un problema de optimización de parámetros. Utiliza probabilidades de cruzamiento y mutación adaptivas. La selección usa elitismo, y como operador usa selección por ruleta. Para el cruzamiento, en las primeras generaciones individuos al azar son apareados. Posteriormente, solo se aparean los individuos que sean suficientemente diferentes entre si. Para ello se usa distancia de hamming. Para la mutación, la probabilidad de mutaciòn es en funciòn del fitness del individuo y los fitness maximo y promedio de la población, de manera que individuos superiores al promedio experimentan...

Desde una perspectiva actual, (Gonçalves and Resende, 2011) proponen un AG basado en claves aleatorias, hibridizado con una nuevo procedimiento de colocación para el *TDC* sin la restricción de guillotina. La representación usa un alfabeto de números reales aleatorios entre 0 y 1 (claves aleatorias). El cromosoma está compuesto de dos partes: la primera contiene la secuencia de piezas a empacar, mientras que la segunda indica el tipo de procedimiento de colocación usado para colocar cada rectángulo. Los autores justifican una representación en dos etapas debido a la dificultad de una representación directa de patrones de corte, en particular la dificultad de desarrollar operadores de cruzamiento y mutación. Para generar la secuencias de piezas, las *N* piezas del problema se biyectan con los primeros *N* genes del cromosoma, Luego de ser generados, estos son ordenados de manera ascendente, obteniendo así, una secuencia (ver figura 2.10). Para los *N* siguientes genes, un valor menor o igual a 0.5 indica que se debe usar la heurística *BL*. En caso contrario se debe emplear la heurística *LB*.



FIGURA 2.10 *Decodificación del cromosoma en representación de códigos aleatorios*

Para el proceso evolutivo emplean una estrategia elitista, preservando los mejores individuos en una porción de la población, *TOP*, reservada para ellos (ver figura 2.11). Para el cruzamiento utilizan *cruzamiento uniforme* parametrizado, donde uno de los padres es seleccionado del pool *TOP*, y el otro es seleccionado al azar de la población. El aportador de cada alelo se determina lanzando una moneda cargada al mejor individuo, según una probabilidad de cruzamiento, la cual fijan experimentalmente en 0.7. En lugar de emplear un operador de mutación, en cada generación se generan nuevos individuos aleatorios que reemplazan al pool de los peores individuos de la población, *BOT*.



FIGURA 2.11 *Proceso de* *transición entre generaciones consecutivas*

La población inicial no es totalmente aleatoria. Se introducen cuatro cromosomas no aleatorios, cuyas secuencias de piezas están en un orden descendente de sus valores, considerando las siguientes combinaciones heurísticas: *random*, todas *BL*, todas *LB*, y *LB* alternado con *BL*. Resultados experimentales muestran que la inclusión de estos cuatro individuos mejora de manera significativa la calidad de las soluciones obtenidas. Proponen una modificación a la función de *fitness* natural (valor total de las piezas cortadas), la cual incorpora información sobre el potencial de mejoramiento del patrón; así, dos patrones con el mismo valor total, pueden tener diferentes potenciales de mejoramiento, dados por la forma en que las pérdidas están distribuídas; si están concentradas en una zona, será mas probable la inserción de piezas en ella. (ver figura 2.12). Así, definen...Para sus experimento configuran el *GA*, mediante un estudio piloto, de donde obtienen los siguientes ajustes: *TOP*= 25%, *BOT*=15%, *CProb*=0.7. El tamaño de población es referenciado al tamaño del problema, dado por la cantidad de piezas. Así, determinaron como tamaño 15 veces la cantidad de rectángulos de la instancia.



FIGURA 2.12 *Dos patrones con mismo valor total y diferentes potenciales de mejora*

Realizan una implementación paralela, en la que paralelizan la evaluación del fitness, debido a su alto consumo de cómputo. Para ello emplean un esquema distribuido, en el cual los mejores dos individuos son compartidos de manera síncrona, con una frecuencia de migración determinada experimentalmente, al resto de las subpoblaciones. Comparan el algoritmo con otros enfoques, concistentes en una heuristica basada en poblaciones PH, un algoritmo genético GA, GRASP y TABU. Concluyen que el enfoque propuesto es efectivo y robusto comparado con los otros enfoques.

## 2.2. ALGORITMOS GENÉTICOS CELULARES

En esta sección se presentan algunos antecedentes históricos del origen y desarrollo de los algoritmos genéticos celulares. Además se citan algunas evidencias, tanto teóricas como empíricas, que sustentan la noción de que los conceptos celulares permiten mejorar las características de la busqueda realizada por un algoritmo genético, para ciertos escenarios y tipos de problemas.

Los algoritmos evolutivos celulares fueron diseñados inicialmente para funcionar en máquinas masívamente paralelas, En el caso más simple, un solo individuo era asignado a un procesador, y el apareamiento entre individuos estaba restringido al individuo más cercano. (Bethke, 1976) hizo el estudio teórico de un AG en una máquina paralela SIMD, analizando la eficiencia del uso de la capacidad de procesamiento. Concluyó que la máxima eficiencia es obtenida cuando la evaluación de la función objetivo es mucho más costosa que los operadores genéticos, lo cual sucede a menudo.

El primer modelo celular conocido (cGA) es el propuesto por (Robertson, 1987), implementado en un computador CM1. En este modelo todos los operadores genéticos eran ejecutados en paralelo. El principal resultado de este trabajo es que el tiempo de ejecución era independiente del tamaño de la población.

(Muhlenbein et al., 1988) publicaron un trabajo donde un cGA en máquinas masívamente paralelas fue propuesto para el TSP. Incorporaron un paso de búsqueda local para mejorar las soluciones generadas. Así, es considero como el primer cGA híbrido publicado.

El término algoritmo genético celular no fue usado sino hasta 1993, cuando (Whitley, 1993) lo propuso por primera vez en un trabajo donde un modelo de autómata celular era aplicado a un algoritmo genético.

Si bien todos estos cGAs fueron inicialmente diseñados para trabajar en máquinas masívamente paralelas, debido a la rápida pérdida de popularidad sufrida por estas máquinas, el modelo fue adoptado despues para trabajar en máquinas mono- procesador, lo que evidencia que el modelo celular es independiente de la arquitectura sobre la cual está implementado.

Una forma simple para caracterizar la búsqueda realizada por un cGA es usar la presión selectiva, la cual es una medida de la velocidad de difusión de las buenas soluciones a través de la población. Algunos trabajos teóricos comparan los algoritmos de acuerdo a la presión selectiva mostrada, y en algunos casos incluso intentan modelar matemáticamente su comportamiento.

(Sarma and De Jong, 1996) realizaron un estudio teórico sobre la presión selectiva inducida por los cGAs con diferentes operadores de selección, y tamaños y formas de vecindades. Para estudiar el efecto del tamaño de la vecindad en la presión selectiva, propusieron una definición del radio de la vecindad como medida de su tamaño. Mas aún, observaron el mismo efecto al cambiar el tamaño de la población, con lo cual propusieron una nueva medida llamado *ratio*, definida como la relación entre el radio de la vecindad y la población. Descubrieron que el ratio es un factor clave para controlar la presión selectiva del algoritmo. Así, dos algoritmos con diferentes tamaños de población y vecindades, pero con el mismo *ratio* tienen una presión selectiva similar. Finalmente propusieron el uso de una función logística para aproximar la curva de presión selectiva de los cGAs. El modelo propuesto pareció ser un buen enfoque para cGAs con poblaciones cuadradas, pero después fue demostrado que tiene ciertas deficiencias cuando se usan poblaciones rectangulares.

(Sprave, 1999) propuso una descripción unificada de cualquier tipo de EA (algoritmo evolutivo) con ambas, poblaciones estructuradas como no estructuradas, basado en el concepto de hipergrafo. Un hipergrafo es una extensión de un grafo canónico, donde el concepto de arco es generalizado: en lugar de unión entre pares de vértices son uniones de subconjuntos de vértices. Usando el concepto de hipergrafo, Sprave desarrolló un método para estimar la curva de crecimiento de presión selectiva de un GA. Este método está basado en el cálculo del diámetro de la estructura de la población y la probabilidad de la distribución inducida por el operador de selección.

(Gorges-Schleuter, 1999) estudió las curvas de crecimiento para un modelo celular de estrategia evolutiva (ES) con poblaciones estructuradas en formas toroidales o anillo. En su estudio, observó que el modelo celular (tanto el toroidal como anillo) tienen menor presión selectiva que el ES equivalente con población no estructurada. Mas aun, comparando los dos modelos, concluyó que, usando el mismo tamaño de vecindad, estructurar la población en forma de anillo permite una menor presión selectiva que al usar una población toroidal.

(Giacobini et al., 2003) propusieron modelos cuantitativos para estimar el takeover time (el tiempo para colonizar la población mediante copias del mejor individuo bajo los efectos de la selección), para cGAs síncronos y asíncronos con una población estructurada en forma de anillo, y usando una vecindad compuesta por los dos individuos más cercanos al individuo considerado. Este trabajo fué extendido con el fin de encontrar modelos matemáticos precisos para fijar las curvas de presión selectiva de cGAs síncronos y asíncronos. Posteriormente los mismos autores propusieron ciertas recurrencias probabilisticas para modelar el comportamiento de la presión selectiva de cGAs síncronos y asíncronos con poblaciones cuadradas, toroidales y lineales (anillo) para dos esquemas de selección diferentes. Este modelo no es completamente preciso cuando son usados otros esquemas de selección.

(Giacobini et al., 2005) propusieron algunos modelos matemáticos para aproximar las curvas de crecimiento de cGAs con poblaciones donde la topología es definida como un grafo aleatorio, donde la distancia entre dos individuos es en general mucho menor que en el caso de las

(Simoncini et al., 2006) propusieron un nuevo operador de selección para cGAs llamado *selección anisotrópica*, para ajustar la presión selectiva del algoritmo. Esta consiste en permitir la selección de individuos de la vecindad con distintas probabilidades de acuerdo a su posición. De esa manera, los autores promueven la aparición de nichos en la población.

Finalmente en (Dorronsoro and Alba, 2007) fue presentada una ecuación matemática más precisa para modelar las curvas de presión selectiva de cGAs con poblaciones rectangulares y cuadradas. El modelo propuesto fue

En la literatura existen resultados que sugieren, pero no analizan, que la forma de la grilla en la población realmente influye en la calidad de la búsqueda realizada por el algoritmo. Como se mencionó anteriormente el concepto de ratio es relevante porque algoritmos con ratios similares muestran un comportamiento similar en la búsqueda.

(Alba and Troya, 2000) publicaron un estudio cuantitativo de las mejoras obtenidas en la eficiencia de un cGA al usar grillas no cuadradas. En este trabajo, el comportamiento de algunos cGAs con diferentes formas de grilla fue analizado en distintos problemas, concluyendo que el uso de grillas no cuadradas promueve un comportamiento eficiente en los algoritmos. Mas aun, redefinieron el concepto de radio como la dispersión de un conjunto de patrones, la cual es más robusta que la definición previa de Sarma y De Jong. Adicionalmente, proponen cambiar dinámicamente la forma de la población, (ratio dinámico), y así autoajustar el balance entre exploración y explotación.

En esta misma linea, (Dorronsoro and Alba, 2005) desarrollaron un nuevo modelo adaptivo en el cual la forma de la población es modificada automáticamente para regular el balance entre exploración y explotación. Diferentes versiones del nuevo algoritmo adaptivo fueron comparados a los algoritmos con ratio estático, y las superaron a todas ellas en todos los casos.

A continuación se citan algunos importantes trabajos que se centran en el análisis del comportamiento de los cGAs, como el proceso evolutivo de los individuos en la población, o la complejidad del algoritmo de acuerdo a los operadores usados.

(Collins and Jefferson, 1991) caracterizan la diferencia entre los GA no estructurados y los cGAs de acuerdo a ciertos factores, como la diversidad del genotipo y fenotipo, la velocidad de convergencia, o la robustez del algoritmo, concluyendo que el apareamiento local realizado por los cGAs es más apropiado para la evolución artificial. Demuestran que para un problema particular con dos óptimos, un GA no estructurado raramente encuentra ambas soluciones, mientras que cGA generalmente las encuentra. Esto es debido a la lenta difusión de las mejores soluciones producidas por el cGA, la diversidad es mantenida por más tiempo en la población, formando pequeños nichos (grupos de individuos similares), representando diferentes áreas de búsqueda del algoritmo. Este trabajo motivó a otros autores a usar cGAs para encontrar óptimos múltiples para problemas. De este y otros trabajos similares se concluyó que un comportamiento característico de los cGAs es la formación de diversos nichos en la población donde el ciclo reproductivo tiende a promover la especialización de los individuos al interior de ellos. Así, los cGAs mantienen diversas rutas de búsqueda hacia diferentes soluciones, donde cada uno de estos nichos puede ser visto como una ruta de explotación del espacio de búsqueda.

(Manderick and Spiessens, 1991) publicaron un estudio comparativo de la complejidad temporal entre su cGA y un GA secuencial. mostrando que para el cGA esta aumenta linealmente de acuerdo al largo del genotipo. Por el contrario, la complejidad de un GA secuencial aumenta polinomialmente de acuerdo al tamaño de la población multiplicado por el largo del genotipo. En este artículo los autores deducen el npumero esperado de individuos al usar los métodos de selección comunes en cGAs, mostrando que la selección proporcional es la que tiene menor presión selectiva.

(Sarma and De Jong, 1995) compararon varios cGAs usando diferentes esquemas de selección y observaron que dos de las selecciones estudiadas se comportaron de manera distinta aún teniendo presiones selectivas equivalentes, desmintiendo así, la asunción de que presiones selectivas equivalentes implican comportamientos similares de búsqueda.

(Gordon et al., 1994) estudiaron siete cGAs con diferentes vecindades en problemas de optimización discretos y continuos. En su trabajo concluyeron que las vecindades más grandes trabajan mejor con problemas más simples, pero al contrario, con problemas más complejos es mejor el uso de vecindades más pequeñas.

Más recientemente, (Alba et al., 2002) realizaron un estudio comparativo del comportamiento de cGAs con políticas de actualización síncronas y asíncronas. Los resultados obtenidos muestran que los cGAs asíncronos ejercen una mayor presión selectiva que los síncronos, por lo que convergen más rápido, y generalmente, encuentran la solución más pronto que los síncronos en los problemas menos complejos estudiados. Por el contrario, en el caso de los problemas más duros, los cGAs síncronos parecen ser los que ofrecen una mejor eficiencia, ya que los asíncronos se atascan en óptimos locales más frecuentemente.

A pesar de que existen diversas extensiones y mejoras al modelo canónico de cGAs, aqui solo se presentan los relacionados al caso simple. Sin embargo la gihu

Del estado del arte se concluye que, en general, soluciones competitivas son obtenidas al aplicar conocimiento especifico del problema, lo cual es atribuible a su fuerte componente geométrica. Por ejemplo, tanto para métodos exactos como aproximados, se suelen usar *lower-bounds* de alta precisión (respecto al óptimo) provistos por heurísticas o relajaciones del problema en la restricción de las piezas, o en el número de etapas para realizar los cortes... También suelen utilizarse heurísticas orientadas a ciertas clases especiales de patrones, los cuales a menudo exhiben propiedades geométricas que se traducen en buenas soluciones, como por ejemplo los patrones de bloque...A menudo se usan reglas heurísticas basadas en el preordenamiento de las piezas en función de sus características geométricas, como su área, ancho, largo, etc. Por lo que una primera pregunta que podriamos formularnos, es si en un escenario en que no se cuente con estas nociones, en este caso de la geometría del problema, ¿podría la evolución por si sola revelarnos este conocimiento? ¿podrán los conceptos celulares promover la confiabilidad en la extracción de dicho conocimiento?

El problema ha sido ampliamente abordado desde el enfoque evolutivo, en particular mediante AG. Respecto a los distintos enfoques

Por otra parte, se ha reconocido la existencia de problemáticas en el comportamiento de la búsqueda respecto a la convergencia del algoritmo, y han habido diferentes formas de alivianar este problema. El enfoque celular por su parte ha sido aplicado a diversos problemas de manera exitosa. Los problemas quizá más relacionados son TSP y VRP. Aunque para estos problemas se usaron enfoques híbridos, combinandolo con operadores de búsqueda local. Una característica del enfoque celular, es que gracias a permite facilmente modificar el comportamiento en la búsqueda del algoritmo,

Hasta ahora, la estructuración de poblaciones usando un enfoque celular, no se ha usado como un mecanismo para mejorar la búsqueda del algoritmo genético para este problema...

# MATERIALES Y MÉTODOS

## 3.1 ALGORITMO GENÉTICO CELULAR

## 3.2 MODELAMIENTO DEL PROBLEMA

### 3.2.1 Esquema general

Al utilizar un enfoque evolutivo, la personalización de la metaheurística se centra en dos tareas, por una parte, decidir la representación genética de una solución candidata para el problema, y por otra, decidir los operadores genéticos apropiados para el problema en cuestión. Para la representación genética se utiliza el esquema genotipo-fenotipo, el cual involucra la definición del genotipo y el mapeo al fenotipo mediante una función constructora. El uso de una codificación especial para las soluciones no es un requisito, en lugar de ello la búsqueda genética puede llevarse a cabo directamente en el espacio fenotípico (tómese como ejemplo el AG de Bortfeldt 2008). El propósito del esquema genotipo-fenotipo es proveer un marco de trabajo para el diseño de algoritmos evolutivos, el cual permite al proceso evolutivo abstraerse del problema particular que se está abordando, enfocándose en aspectos propios de la evolución, como la diversidad genética o la presión de selección.

Los enfoques metaheurísticos que abordan problemas de *C&P*, a menudo utilizan heurísticas de colocación como algoritmos decodificadores de una solución particular; Las heurísticas de colocación son algoritmos que reciben una secuencia de piezas y sistemáticamente van determinando su disposición en la placa. La idea es que la metaheurística se encargue de realizar la búsqueda en el espacio de soluciones, mientras que la heurística de colocación se encarga de realizar la traducción a una solución válida para el problema. El esquema general es descrito en la ilustración xx, se parte de un espacio genotípico, donde tiene lugar el proceso evolutivo. Cada genotipo es mapeado mediante la función constructora a un fenotipo, consistente en una secuencia de piezas. El fenotipo es traducido a un patrón guillotinable mediante una heurística de colocación. El patrón finalmente es evaluado por una función objetivo para obtener el fitness de la solución candidata.



FIGURA 3.1 *Esquema general: Función constructora + heurística colocación*

### 3.2.2 Representación

La representación utilizada está basada en la propuesta por A. Flores 2012, la cual utiliza una codificación binaria para representar todas las posibles secuencias de piezas[[2]](#footnote-2). Para ello, el cromosoma se divide en dos partes: un cromosoma de piezas, que decide cuáles piezas estarán presentes en la secuencia, y un cromosoma de ordenamiento, el cual codifica qué orden tienen las piezas en la secuencia.

Para el cromosoma de piezas se hace una biyección entre cada una de las copias de los distintos tipos de pieza y cada gen del cromosoma, de esa manera cada posición referencia a una copia. De esto se deduce que el largo del cromosoma de piezas es equivalente a la suma de todas las copias de cada tipo de pieza. Para cada gen un valor 1 indica que la copia de la pieza está presente en la secuencia, y un valor 0 indica que no lo está. Considérese el ejemplo de la figura x.x.a, con una demanda de 6 tipos de piezas. Para cada tipo se definen sus dimensiones y su límite superior. Así, la pieza de tipo 1 tiene 3 copias, la de tipo 2 tiene 2 copias, la de tipo 3 tiene 1 copia, etc.

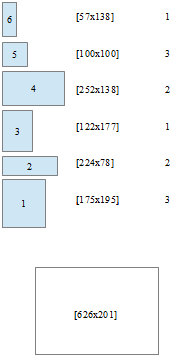
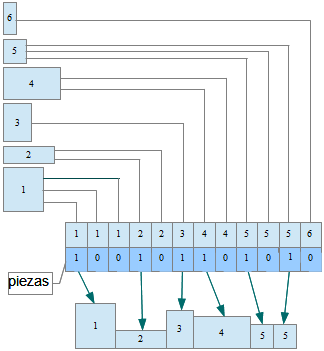


FIGURA 3.2.b *Cromosoma de piezas e interpretación*

FIGURA 3.2.a *Instancia de ejemplo problema*

Las piezas son serializadas en un arreglo que mantiene cada una de las copias, y que es usado como referencia para interpretar cada gen del cromosoma. En la figura x.x.b se ilustra la interpretación de un cromosoma de ejemplo, según el cual: la 1a copia de la pieza de tipo 1 está presente en la secuencia, la 2a y 3a copias no lo están, la 1a copia de tipo 2 si lo está, etc.

Los genes de ordenamiento determinan el orden de las piezas mediante una lectura del cromosoma de piezas. La idea es que durante la lectura, cada vez que se encuentre una pieza marcada como presente (gen=1), esta sea agregada a la secuencia ordenada. Para que la lectura genere diversidad de ordenamientos se definen dos parámetros: sentido y salto. El sentido establece si la lectura se realiza desde la primera posición hasta la última, o viceversa. El salto establece cada cuántas posiciones debe realizarse una lectura. Para el sentido solo se necesita 1 bit ya que tiene dos valores posibles: 0 para un sentido del inicio al final, y 1 para un sentido opuesto. Para el salto, si el largo del cromosoma de piezas es n, se necesitan ld(n) bits para representar todos los saltos posibles que se pueden generar.

El cromosoma completo es el que se muestra en la figura x.x.a. El largo total para codificar una solución es: n+ld(n)+1, donde n corresponde a la cantidad total de piezas, ld(n) al salto en la lectura del cromosoma de piezas, y 1 al sentido de la lectura, la cual se realiza mediante saltos sucesivos en forma cíclica hasta haber visitado todas las piezas. Es por esto que el cromosoma se interpreta conceptualmente en una disposición toroidal, obteniéndose una ruleta de piezas como la que se muestra en la figura x.x.b. A diferencia del resto de los genes, los cuales tienen una interpretación directa, para obtener el valor del salto se hace una conversión de binario a entero[[3]](#footnote-3). La interpretación completa del genotipo se describe mediante el cromosoma del ejemplo: Las 1as copias de las piezas de tipo 1, 2, 3, 4, y la 1a y 3a copia de la pieza de tipo 5 están presentes en la secuencia. La lectura de la ruleta de piezas debe ser en el sentido de las manecillas del reloj y en saltos sucesivos de 5 posiciones.

|  |
| --- |
| FIGURA 3.4.a *Cromosoma completo y su interpretación* |
| FIGURA 3.4.b *Interpretación conceptual como* "*Ruleta de piezas"* |

### 3.2.3 Función Constructora

Para decodificar el string binario en una secuencia de piezas, lo primero que realiza la función constructora es descomponer el cromosoma de acuerdo a los desplazamientos asociados a cada parte; el sentido corresponde a la 1a posición, el salto, de la posición 1 hasta 1+ld(n), y las piezas desde la posición 1+ld(n) hasta 1+ld(n)+n. De esta manera se obtiene el cromosoma de piezas y los parámetros necesarios para a continuación proceder con su lectura.

En la figura x.x se muestra el pseudocódigo de la función constructora, que recibe como parámetro el genotipo, además existe una estructura global “piezas” que corresponde a un arreglo que almacena las distintas piezas del problema. Las líneas 1-6 corresponden a la obtención de los distintos parámetros e inicialización de variables para realizar la lectura. Las líneas 7-20 corresponden a los ciclos de lectura, al final de cada ciclo interno se verifica si se ha cumplido un ciclo en la sucesión de saltos, de ser así, se desliza la sucesión de saltos en una posición, este mecanismo permite visitar las piezas que no han sido visitadas hasta el ciclo de saltos actual.

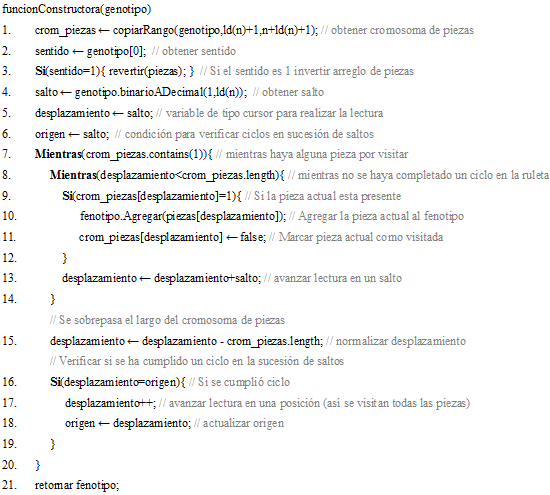


FIGURA 3.5 *Pseudocódigo función constructora*

Para ver el funcionamiento de la función constructora se incluye la traza para la solución candidata de ejemplo. Inicialmente se realiza la descomposición del cromosoma, extrayendo la ruleta de piezas, el sentido y salto. Como el sentido es 0 la lectura se hace en el sentido de las manecillas del reloj, la conversión a entero del salto corresponde a 5, así se fija el desplazamiento y el origen en 5. Luego comienza el ciclo de lectura, la 1a pieza leída es la 3, la cual está presente, por lo tanto se agrega a la secuencia ordenada, el desplazamiento se incrementa en un salto (5+5=10), este desplazamiento corresponde a la pieza 5, la cual está presente, por lo tanto se agrega a la secuencia. Al incrementar el desplazamiento (10+5=15), se sobrepasa el largo del cromosoma de piezas (12), por lo que es normalizado (15-12=3), este desplazamiento corresponde a la pieza 2, la cual está presente, por lo tanto se agrega a la secuencia, y así sucesivamente hasta haber visitado todas las piezas. En este caso, al cumplirse un ciclo en la sucesión de saltos todas las piezas ya han sido visitadas, por lo que el ciclo de lectura termina y no es necesario deslizar los saltos.

|  |  |  |
| --- | --- | --- |
|  | | |
|  | | |
|  |  |  |

FIGURA 3.6 *Traza función constructora; obtención de ruleta de piezas, parámetros de lectura y tres primeras iteraciones de lectura de piezas*

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
|  |  |

FIGURA 3.7 *Traza función constructora, resto de iteraciones de lectura de piezas*

## 3.2.4. Heurística de colocación

Una vez obtenida la secuencia de piezas, la heurística de colocación realiza la traducción a un patrón guillotinable. La heurística utilizada está basada en la *Combinación Heurística VH* (Romero 2003), la cual emplea un enfoque *bottom-up*, donde cualquier patrón que satisfaga la condición de guillotina puede ser obtenido mediante construcción horizontal o vertical de rectángulos. Una combinación vertical consiste en ubicar una pieza arriba de un patrón existente, creando un nuevo patrón de corte, tal como muestra la Ilustración x.x.a. Una combinación horizontal consiste en ubicar una pieza a la derecha de un patrón existente, creando un nuevo patrón de corte, tal como muestra la Ilustración x.x.b.

|  |  |
| --- | --- |
| FIGURA 3.8.a *Combinación horizontal de las piezas A y B* | FIGURA 3.8.b *Combinación horizontal de las piezas A y B* |

Como se observa en las figuras, cada combinación tiene asociada una pérdida interna, correspondiente al área rectangular que contiene las piezas y que queda inutilizada. A su vez, el patrón resultante también tiene asociada una pérdida externa, correspondiente al área de la placa que queda inutilizada. En este trabajo se extiende el concepto de combinación vertical-horizontal, de manera que las pérdidas internas son pérdidas, también verticales u horizontales, asociadas a una pieza en particular. Del mismo modo, las pérdidas externas son pérdidas verticales u horizontales asociadas al patrón de corte. Bajo este concepto, en la figura x.x.a la pérdida interna es una pérdida horizontal asociada a la pieza B, la cual no tiene pérdida vertical asociada. De manera similar, en la figura x.x.b la pérdida interna es una pérdida vertical asociada a la pieza B, la cual no tiene pérdida horizontal asociada.

En la figura x.x se muestra el pseudocódigo de la heurística de colocación, que recibe como parámetro el fenotipo, (dado por una secuencia de piezas), y retorna un layout (consistente en un patrón guillotinable). Además existe una estructura global *“listaPerdidasInternas”*, que almacena las pérdidas asociadas a cada pieza. La heurística recorre la secuencia del inicio hasta el final, y para cada pieza evalúa cada uno de los siguientes casos:

* Si la placa está vacía intenta ubicar la pieza actual en el origen. Como esta acción genera pérdida externa se crean las pérdida externas.
* Si la placa no está vacía :
  + busca (en *listaPerdidasInternas*) una pérdida interna donde quepa la pieza. Esta búsqueda puede ser de acuerdo a algún criterio. En este trabajo se ha usado el criterio first-fit[[4]](#footnote-4). Luego se actualizan las pérdidas internas.
  + Si no es posible insertar la pieza en alguna pérdida interna, se busca una pérdida externa donde quepa la pieza. Luego se actualizan las pérdidas externas.
  + Si no es posible insertar la pieza en alguna pérdida externa, se descarta y se lee la pieza siguiente en la secuencia.

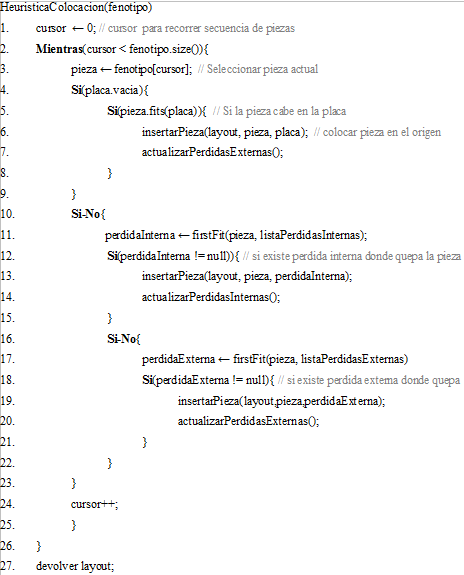


Figura 3.9 *Pseudocódigo heurística colocación*

Para ver el funcionamiento de la heurística se incluye la traza para la solución candidata de ejemplo. Inicialmente, como la placa está vacía, se inserta la pieza 3 en el origen, produciendo una pérdida externa vertical y horizontal asociadas al patrón de corte (figura x.x.2). La siguiente pieza en la secuencia es la 5, como la placa no está vacía y no existen pérdidas internas, se buscan pérdidas externas capaces de contener a la pieza 5, la cual es insertada en la pérdida externa vertical, generando una pérdida interna horizontal asociada a ella, además se reduce la pérdida externa vertical existente (figura x.x.3). La siguiente pieza en la secuencia es la 2, la cual no cabe en ninguna pérdida interna, por lo que es ubicada en la pérdida externa vertical, generando una pérdida interna horizontal asociada a la pieza 3, y una pérdida interna vertical asociada a la pieza 2. Se actualizan las pérdidas externas (figura x.x.4). La siguiente pieza en la secuencia es la 5, la cual no cabe ni en la pérdida horizontal asociada a la pieza 5 ni en la pérdida vertical asociada a la pieza 2, sin embargo cabe en la pérdida horizontal asociada a la pieza 3. Se actualizan las pérdidas internas; se elimina la pérdida utilizada y se crea una nueva pérdida vertical, asociada a la pieza insertada, y así sucesivamente hasta haber leído todas las piezas.

|  |
| --- |
|  |
|  |
|  |
|  |
|  |
|  |

FIGURA 1.10 *Traza heurística colocación 1a parte*

|  |
| --- |
|  |
|  |

FIGURA 3.11 *Traza heurística colocación 2a parte*

### 3.2.5 Función objetivo

Finalmente el patrón guillotinable es evaluado por la función objetivo para obtener el fitness de la solución candidata, el cual representa su competitividad para el problema en cuestión. Como el objetivo es maximizar el área total utilizada en la placa, el fitness está dado por la suma de las áreas de las piezas presentes en el patrón.

...

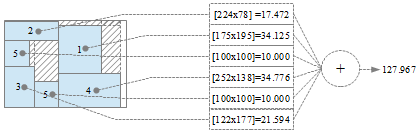


FIGURA 3.12 *Función objetivo y fitness de la solución candidata*

### 3.2.6 Operadores genéticos

Los operadores genéticos constituyen el motor del ciclo reproductivo y así el núcleo evolutivo del AG. Inspirados en procesos biológicos como la selección natural, la reproducción sexual, y la mutación, a continuación se explicarán los operadores genéticos más comunes para una representación binaria, la cual ha sido utilizada en este modelamiento.

**6.1 Selección**

Parte de la evolución está determinada por la selección natural de los individuos en su adaptación al entorno. Inevitablemente, algunos individuos son mejores que otros, éstos son más propensos a sobrevivir, aprender y difundir su material genético. La selección es una función de la aptitud, es decir, qué tan bien compite cada individuo en su entorno. La intensidad que tenga el mecanismo de selección para discriminar cuán apto es un individuo en relación al resto, determinará cuál es el tiempo necesario para que el individuo más sobresaliente se difunda por la población completa, esto se denomina presión de selección, y cada operador posee una presión característica.

**6.1.1 Selección por ranking:** Ordena la población basándose en el fitness y asigna probabilidades de selección de acuerdo a su ranking. El mapeo desde la posición en el ranking a la probabilidad de selección es arbitrario, y puede hacerse de varias formas. a menudo se usa un mapeo lineal o exponencial. En el caso lineal la probabilidad de seleccionar un individuo i es:



Figura 3.13 *Probabilidad de seleccionar individuo i mediante ranking lineal*

donde n es el tamaño del ranking y j es la posición del individuo i en el ranking. En general un mapeo lineal limita la presión de selección, mientras que un mapeo exponencial pone mayor énfasis en seleccionar individuos cuyo fitness supera la media.

**6.1.2 Selección por ruleta:** .Se simula el giro de una ruleta, la cual es particionada en proporción a los fitness de cada individuo. La ruleta se gira n veces para seleccionar n individuos. La probabilidad de seleccionar un individuo es proporcional a su fitness:

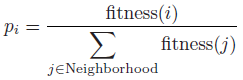


Figura 3.14 *Probabilidad de seleccionar individuo i mediante ruleta*

Para simular el giro de la ruleta se asume algún orden en la población (ranking o aleatorio). Cada vez que se gira la ruleta se genera un valor aleatorio r uniforme en [0,1], luego para cada individuo en la población si su probabilidad de selección pi supera a r entonces es seleccionado.

**6.1.3 Selección por torneo:** Un grupo k de individuos, típicamente dos, son seleccionados aleatoriamente y el mejor es considerado como el padre. A diferencia de los operadores anteriores, la selección por torneo no requiere conocimiento global de la población. En lugar de ello, solo depende de una relación de orden que permite rankear cualquier par (o más) individuos. Similar a un esquema de ranking, utiliza un fitness relativo en vez de absoluto. La probabilidad de que un individuo sea seleccionado depende de:

* Su ranking en la población. Efectivamente esto es estimado sin necesidad de ordenar la población completa
* El tamaño k del torneo. Mientras más individuos considere, mayor es la oportunidad de que contenga individuos superiores a la media
* La probabilidad p de que el individuo más apto del torneo sea seleccionado
* Si los individuos son escogidos con o sin reemplazo

La selección

**6.2 Cruzamiento**

Tal como en la reproducción sexual, el cruzamiento considera el apareamiento entre dos individuos[[5]](#footnote-5), produciendo crías que contienen una combinación de la información de sus padres. La importancia del cruzamiento es su capacidad de reunir propiedades de soluciones candidatas que pueden estar localizadas en regiones distantes entre sí. Esto favorece la exploración, ya que permite descubrir áreas prometedoras en el espacio de búsqueda, mediante grandes saltos a áreas que se encuentran entre dos padres.

**6.2.1 Cruzamiento de 1 punto:** Se escoge una posición aleatoria en el cromosoma, luego se dividen ambos padres en ese punto y se crean dos hijos intercambiando las colas.

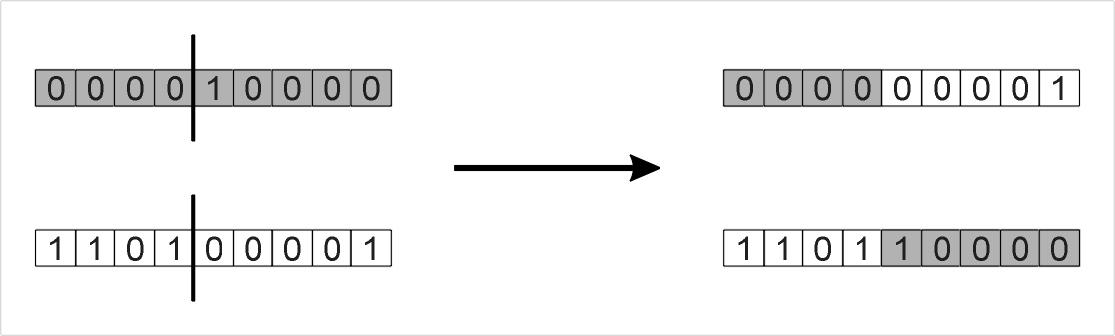


Figura 3.15 *Cruzamiento de 1 punto*

**6.2.1 Cruzamiento de n puntos:** Es la generalización del cruzamiento de 1 punto, en la cual el genotipo es fragmentado en más de dos segmentos de genes continuos, luego los hijos son creados tomando segmentos alternados de los padres.

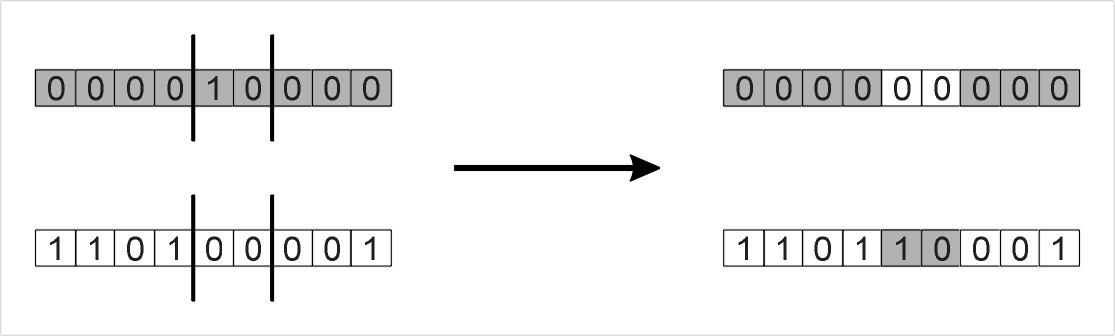


Figura 3.16 *Cruzamiento de n puntos; n=2*

Una característica importante del cruzamiento de n puntos es el sesgo posicional, ya que tiende a mantener juntos genes que están cercanos el uno del otro en el genotipo. En general a medida que la cantidad de puntos de cruce aumenta, este sesgo tiende a disminuir.

**6.2.3 Cruzamiento uniforme:** Cada gen es tratado de manera independiente, y por cada uno se escoge aleatoriamente de qué padre debe ser heredado

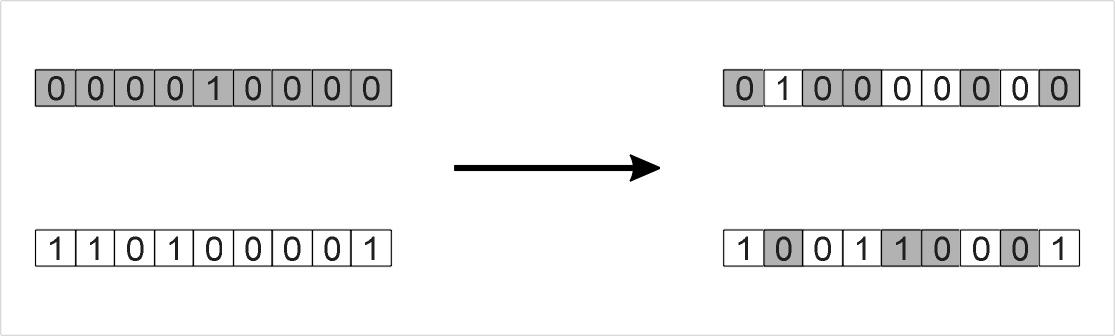


Figura 3.17 *Cruzamiento uniforme*

A diferencia del cruzamiento de n puntos, el cruzamiento uniforme no exhibe sesgo posicional, sin embargo tiene una fuerte tendencia a transmitir el 50% de los genes de cada padre y en contra de transmitir un mayor número de genes coadaptados. Esto es conocido como sesgo distribucional.

**6.3 Mutación**

Así como en el caso biológico, los individuos pueden mutar ocasionalmente experimentando pequeños cambios en su información genética. La mutación permite la explotación de áreas prometedoras, realizando pequeños movimientos dentro de dicha región. También es un mecanismo que ayuda a mantener cierto grado de diversidad en la población (evitando atascarse en óptimos locales).

**Mutación binaria:** El operador más comúnmente usado para representaciones binarias considera cada gen separadamente y permite a cada bit cambiar con una pequeña probabilidad pm. Luego para una codificación de largo L, en promedio Lpm

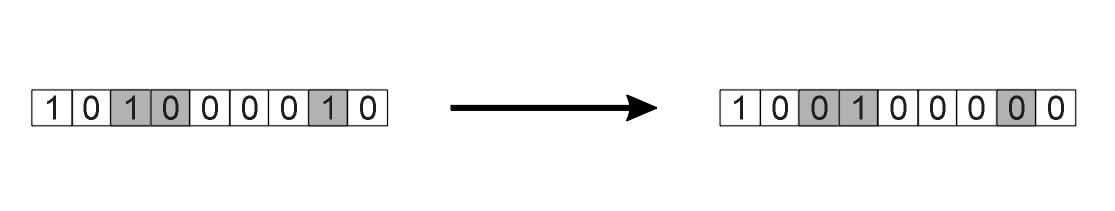


Figura 3.18 *Mutación binaria*

lalalala

lalalala

## 3.2 DATOS DE PRUEBA

## 3.4 CONFIGURACIÓN DE ALGORITMOS

### 3.4.1 Motivación

Muchos algoritmos de alto rendimiento poseen parámetros cuyos ajustes controlan importantes aspectos de su comportamiento. Este es el caso de los procedimientos heurísticos, usados para resolver problemas computacionalmente difíciles. Hay muchos reconocimientos en la literatura de que la búsqueda de configuraciones que optimicen el rendimiento de estos algoritmos a menudo requiere un esfuerzo considerable. En muchos casos, esta tediosa tarea es realizada manualmente de manera ad-hoc, por lo tanto su automatización es de gran importancia práctica en varios contextos:

* **Desarrollo de algoritmos complejos:** El ajuste de parámetros de un algoritmo heurístico es una tarea muy laboriosa, y de hecho puede consumir una gran parte del tiempo del desarrollo general. El uso de métodos automatizados de configuración de algoritmos puede conducir a importantes ahorros de tiempo y potencialmente lograr mejores resultados que métodos manuales ad-hoc
* **Estudios empíricos, evaluaciones y comparaciones de algoritmos**: Una cuestión central en la comparación de algoritmos heurísticos es si un algoritmo supera a otro porque es fundamentalmente superior, o porque sus desarrolladores fueron más exitosos optimizando sus parámetros. Los métodos automatizados de configuración pueden mitigar este problema de comparaciones injustas y así facilitar estudios comparativos más significativos
* **Uso práctico de los algoritmos:** La habilidad de algoritmos heurísticos complejos para resolver instancias grandes y difíciles a menudo depende en forma crítica de un adecuado ajuste de parámetros. Los usuarios finales usualmente tienen poco o ningún conocimiento del impacto del ajuste de los parámetros en el rendimiento del algoritmo, y así usan simplemente el ajuste por defecto. Aún cuando haya sido cuidadosamente optimizado con un benchmark estándar, tal configuración puede no funcionar bien en la instancias particulares encontradas por un usuario. Los métodos automatizados de configuración pueden ser usados para mejorar el rendimiento con principios y de manera conveniente

El problema de la configuración de algoritmos puede ser enunciado informalmente como: Dado un algoritmo, un conjunto de parámetros para el algoritmo y un conjunto de datos de entrada, encontrar los valores para los parámetros bajo los cuales el algoritmo logre el mejor rendimiento posible con los datos de entrada. Para diferenciar los algoritmos cuyo rendimiento debe ser optimizado de los algoritmos utilizados para esa tarea, los primeros son denominados algoritmos objetivo, mientras que los últimos son llamados configuradores.

La figura x.x ilustra el problema de configuración, un escenario de configuración incluye un algoritmo y una colección de instancias del problema. Un configurador ejecuta el algoritmo objetivo con un determinado ajuste de parámetros sobre alguna o todas las instancias, recibe información sobre el rendimiento de estas ejecuciones, y utiliza esta información para decidir las siguientes configuraciones de parámetros evaluar.



Figura 3.18 *Problema de configuración de algoritmos*

Teniendo claros los elementos que intervienen en el problema de configuración, resulta fácil identificar cada uno de ellos. En nuestro caso tenemos varios algoritmos objetivo correspondientes a cada uno de los AG propuestos para la comparación, por lo que debe resolverse un problema de configuración por cada uno de ellos. Por su parte, las instancias corresponden al conjunto de instancias seleccionado, las cuales son descritas en la sección x.x. El configurador corresponde a la herramienta de optimización de parámetros a utilizar, la cual será descrita a continuación.

### 3.4.2 ParamILS

ParamILS es un framework para la optimización de parámetros, el cual utiliza la metaheurística basada en trayectoria, Iterated Local Search (de ahí su nombre). Funciona para cualquier algoritmo cuyos parámetros puedan ser discretizados. ParamILS busca a través del espacio de posibles configuraciones de parámetros, evaluando configuraciones mediante la ejecución del algoritmo objetivo en un conjunto de instancias de prueba. Para ello se debe definir:

* Un algoritmo ejecutable *A* (algoritmo objetivo)
* Todos los parámetros y sus valores posibles
* Un conjunto de instancias de prueba *S*

Existen distintas maneras de medir el desempeño del algoritmo objetivo. Por ejemplo, puede interesar minimizar el costo de los recursos computacionales consumidos (como tiempo de ejecución, memoria, ancho de banda), o maximizar la calidad de la solución encontrada. Para tal efecto, ParamILS permite elegir de una variedad de objetivos de optimización, desde minimizar el tiempo medio de ejecución hasta maximizar la mediana de las calidades de solución. Así, ParamILS ejecuta el algoritmo A sobre una muestra de instancias de S, en busca de la configuración que obtenga el mejor rendimiento general. Para un… referencias

lalalalala

aqui agregar marco teórico

### 3.4.3 Plan de configuración y diseño de escenarios de sintonización

Distintos algoritmos objetivo, correspondientes a cada AG propuesto, deben ser configurados para optimizar su desempeño sobre el conjunto de instancias seleccionado. Al mismo tiempo se desea realizar comparaciones entre los desempeños de cada algoritmo. Para mantener la consistencia en dichas comparaciones se define un plan de configuración, el cual prioriza los algoritmos con una menor cantidad de parámetros, comunes para todos ellos. De este modo, la configuración de un algoritmo posterior asumirá el ajuste de los parámetros previamente obtenidos, y solo requerirá la obtención de parámetros adicionales, específicos para dicho algoritmo.

La figura x.x muestra los valores posibles que se consideraron para cada parámetro, un ticket indica que el parámetro es obtenido en una configuración previa y un guión indica que el parámetro no aplica para ese algoritmo. Dado que los AG no estructurados poseen parámetros comunes a todos, uno de ellos, el AG-generacional, será el primero en ser configurado, de cuyo resultado se obtendrá la totalidad de parámetros para el AG steady-state y parte de los parámetros para los AG estructurados. En particular, para el AG-distribuido, una previa obtención del tamaño de la población permite fijar el tamaño de las islas, el cual se obtiene dividiendo el tamaño de la población por la cantidad de islas. En el caso del AG-celular, una previa obtención del tamaño de población permite fijar el área de la cuadrícula bidimensional o enrejado[[6]](#footnote-6), de manera que los valores posibles para este parámetro son las dimensiones del enrejado. Además se añaden las cardinalidades de los espacios de configuración, esto permite estimar los tiempos de espera para configurar cada algoritmo.

|  |  |  |  |
| --- | --- | --- | --- |
| *algoritmo parámetro* | *AG generacional* | *AG distribuido* | *AG celular* |
| *tamaño población* | 50; 100; 150; 200; 250; 300; 350; 400; 450; 500 | ─ | ─ |
| *probabilidad cruzamiento* | 0,6; 0,7; 0,8; 0,9; 1,0 | ✔ | ✔ |
| *probabilidad mutación* | 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; 0,9; 1,0 | ✔ | ✔ |
| *operador selección 1* | torneo-binario; ruleta; ranking-lineal | ✔ | ✔ |
| *operador selección 2* | torneo-binario; ruleta; ranking-lineal | ✔ | ✔ |
| *operador cruzamiento* | 1-punto; 2-puntos; uniforme; probabilístico | ✔ | ✔ |
| *operador mutación* | por-bit | ✔ | ✔ |
| *cantidad islas* | ─ | 4; 5; 6; 7; 8; 9; 10 | ─ |
| *tamaño islas* | ─ | ✔ | ─ |
| *frecuencia migración* | ─ | 5000; 10000; 15000;  20000; 25000 | ─ |
| *enrejado* | ─ | ─ | 20x20; 10x40; 5x80; 4x100; 2x200; 1x400 |
| *vecindad* | ─ | ─ | Linear5; Compact9; Compact13 |
| *política actualización* | ─ | ─ | ls; frs; nrs; uc; ss; synchronous |
| *# total configuraciones* | 27000 | 35 | 108 |

Tabla 3.1 *Valores posibles para cada parámetro*

*ParamILS* no restringe los valores de los parámetros a dominios únicamente numéricos, por el contrario, asume que todos los parámetros son categóricos. Gracias a esta característica se podria definir por ejemplo un parámetro para seleccionar entre heurísticas de búsqueda. Para los algoritmos objetivo se puede definir varios parámetros categóricos, como los distintos operadores genéticos, y parámetros específicos del AG-celular. Se aprovechó el hecho de que varios de estos se encuentran implementados en la plataforma evolutiva utilizada (*JCell*), y fueron incluidos en el escenario de sintonización.

*ParamILS* obtiene desde un archivo toda la información necesaria para realizar la configuración de parámetros, este archivo es llamado *escenario de* sintonización. A continuación se especifican los principales parámetros de un *escenario de sintonización,* y cómo fueron definidos para la configuración de los algoritmos objetivo:

|  |  |
| --- | --- |
| ***Parámetro*** | ***Descripción*** |
| ***algo*** | Un algoritmo ejecutable o una llamada a un script de envoltura en torno al algoritmo, que cumpla con el formato de entrada/salida de *ParamILS*. Para este trabajo se desarrolló un script de envoltura, el cual entre otras cosas realiza la normalización de la mejor solución encontrada por el algoritmo respecto al valor óptimo de la instancia respectiva. |
| ***run\_obj*** | Un escalar que cuantifique cuán buena es una ejecución del algoritmo. Algunos ejemplos son: tiempo de ejecución, aproximación de la solución, etc. Para efectos de este estudio se prioriza optimizar la calidad de la solución, luego se usó una expresión que mide la aproximación de la mejor solución al óptimo, *qual=óptimo/best* especificando para cada instancia su valor óptimo correspondiente. |
| ***overall\_obj*** | Una medida de resumen que permita describir la tendencia en la calidad de varias ejecuciones del algoritmo. Algunos ejemplos son: promedio, mediana, q90, etc. En este trabajo se utiliza el promedio |
| ***cutoff\_time*** | El tiempo transcurrido tras el cual una ejecución del algoritmo será terminada sin éxito. (Ej: No alcanzar el óptimo). El *cutoff\_time* fue ajustado empíricamente en 260s, tiempo en el cual el AG-generacional efectúa 150.000 evaluaciones sin encontrar el óptimo |
| ***cutoff\_length*** | La longitud tras la cual una ejecución del algoritmo será terminada sin éxito. Esta longitud puede ser definida como la cantidad de decisiones en un árbol de búsqueda o la cantidad de flips en un algoritmo SLS. Para los algoritmos a evaluar se ha definido como la cantidad de evaluaciones de la función objetivo, la cual se ajustó empíricamente en 150.000 evaluaciones, longitud tras la cual los AG en general no mejoraban la mejor solución encontrada. |
| ***tuner\_timeout*** | Tiempo de espera del configurador, tras el cual comienza la validación de la mejor configuración de parámetros encontrada. Los tiempos de espera fueron estimados para cada algoritmo en base al *cutoff\_time* y a los tamaños de sus espacios de configuración, de manera de alcanzar a realizar como mínimo 10 ejecuciones por instancia en el caso pesimista de no encontrar el óptimo en ninguna ejecución. |
| ***instance\_file & test\_instance\_file*** | *ParamILS* utiliza un conjunto de instancias de entrenamiento y un conjunto de instancias de prueba, para validar la configuración obtenida en el entrenamiento. Para ello, el conjunto de 49 instancias se particionó en 50% para entrenamiento y el resto para prueba. Para que la distribución fuera uniforme, las instancias fueron repartidas equitativamente en cada conjunto. |

Tabla 3.2 *Parámetros escenario sintonización paramILS*

La figura x.x muestra el escenario de sintonización para el AG-generacional

alalala

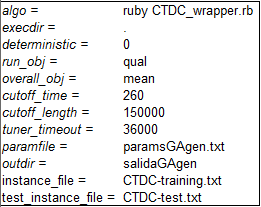


figura x.x.- Ejemplo de escenario de sintonización para el *AG-Generacional*

|  |  |  |  |
| --- | --- | --- | --- |
| *algoritmo parámetro* | *AG generacional* | *AG distribuido* | *AG celular* |
| *tamaño población* | 50; 100; 150; 200; 250; 300; 350; **400**; 450; 500 | ─ | ─ |
| *probabilidad cruzamiento* | 0,6; 0,7; 0,8; 0,9; **1,0** | ✔ | ✔ |
| *probabilidad mutación* | 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; **0,9**; 1,0 | ✔ | ✔ |
| *operador selección 1* | torneo-binario; **ruleta**; ranking-lineal | ✔ | ✔ |
| *operador selección 2* | torneo-binario; **ruleta**; ranking-lineal | ✔ | ✔ |
| *operador cruzamiento* | **1-punto**; 2-puntos; uniforme; probabilístico | ✔ | ✔ |
| *operador mutación* | **por-bit** | ✔ | ✔ |
| *cantidad islas* | ─ | 4; **5**; 6; 7; 8; 9; 10 | ─ |
| *tamaño islas* | ─ | ✔ | ─ |
| *frecuencia migración* | ─ | 5000; 10000; **15000**;  20000; 25000 | ─ |
| *enrejado* | ─ | ─ | 20x20; 10x40; 5x80; 4x100 ;2x200; **1x400** |
| *vecindad* | ─ | ─ | Linear5;**Compact9**; Compact13 |
| *política actualización* | ─ | ─ | ls; frs; nrs; **uc**; ss; synchronous |

Tabla 3.3 *Valores obtenidos por ParamILS para cada algoritmo objetivo*

## 3.4 DISEÑO EXPERIMENTAL

## 3.5 HARDWARE Y SOFTWARE UTILIZADO

# RESULTADOS

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Distribución del *fitness* en la población |  |  |  |  |  |  |  |  |
| Generación | 50 | 150 | 250 | 350 | 450 | 550 | 650 | 750 |
| Mean | 772758 | 786202 | 794762 | 799856 | 801536 | 802506 | 804233 | 805271 |
| Best | 793482 | 800684 | 801656 | 801656 | 801656 | 805106 | 805106 | 809838 |

# ANÁLISIS DE RESULTADOS

# CONCLUSIONES

# REFERENCIAS BIBLIOGRÁFICAS

1. Corresponde al mínimo entre distintas soluciones candidatas. Ej: la solución de la ecuación recursiva del problema de programación dinámica formulado por Gilmore y Gomory (1966), la solución óptima para el problema de programación dinámica del UTDC, entre otras. [↑](#footnote-ref-1)
2. Es fácilmente demostrable que la representación propuesta no es capaz de generar todas las secuencias posibles. Sin embargo se asume que es capaz de generar una diversidad suficiente [↑](#footnote-ref-2)
3. De acuerdo a A. E Eiben 2003 p. 40-41, debido a la significación numérica de bits en una codificación binaria, una codificación de Gray puede ser más apropiada para representar el salto. [↑](#footnote-ref-3)
4. Retorna la primera pérdida donde quepa la pieza. Nótese que dependiendo del resultado de la búsqueda (pérdida vertical u horizontal), la inserción de la pieza producirá una combinación vertical u horizontal respectivamente. [↑](#footnote-ref-4)
5. En un contexto computacional, donde no existen las limitaciones biológicas [↑](#footnote-ref-5)
6. Estructura geométrica típica en AG celulares. Sin embargo existen otras... [↑](#footnote-ref-6)