

UNIVERSIDAD DE GRANADA
E.T.S.I. INFORMÁTICA Y TELECOMUNICACIÓN



Departamento de Ciencias de la
Computación e Inteligencia Artificial

Metaheurísticas

<http://sci2s.ugr.es/graduateCourses/Metaheuristics>
<https://decsai.ugr.es>

Guión de Prácticas

Práctica 2.b:
Búsquedas Multiarranque para el
Problema de la Selección de Características

Curso 2015-2016

Tercer Curso del Grado en Ingeniería Informática

Práctica 2.b

Búsquedas Multiarranque para el Problema de la Selección de Características

1. Objetivos

El objetivo de esta práctica es estudiar el funcionamiento de los siguientes algoritmos: Búsqueda Multiarranque Básica (BMB), GRASP y Búsqueda Local Reiterada (ILS). Para ello, se requerirá que el alumno adapte estos métodos para resolver el problema de la selección de características (SC) descrito en las transparencias del Seminario 2.a y que compare los resultados obtenidos con los proporcionados por el clasificador 3-NN generado considerando todas las características disponibles y con el clasificador 3-NN obtenido empleando las características seleccionadas por el método *greedy Sequential Forward Selection* (SFS) en una serie de casos del problema.

La práctica se evalúa sobre un total de **1,5 puntos**, distribuidos de la siguiente forma: BMB (0,3 puntos), GRASP (0,7 puntos) e ILS (0,5 puntos).

La fecha límite de entrega será el **Lunes 9 de Mayo de 2016** antes de las 23:59 horas. La entrega de la práctica se realizará por internet a través del acceso identificado de la web del departamento de CCIA (<https://decsai.ugr.es>).

2. Trabajo a Realizar

El alumno podrá desarrollar los algoritmos de la práctica siguiendo la modalidad que desee: trabajando con cualquiera de los *frameworks* de metaheurísticas estudiados en el Seminario 1, implementándolos a partir del código C proporcionado en la web de la asignatura o considerando cualquier código disponible en Internet.

Los métodos desarrollados serán ejecutados sobre una serie de casos del problema. Se realizará un estudio comparativo de los resultados obtenidos y se analizará el comportamiento de cada algoritmo en base a dichos resultados. **Este análisis influirá decisivamente en la calificación final de la práctica.**

En las secciones siguientes se describen los aspectos relacionados con cada algoritmo a desarrollar y las tablas de resultados a obtener. Los casos del problema, el número de ejecuciones a realizar sobre ellos y los estadísticos de calidad (*Tasa_clas*,

Tasa_red y *Tiempo*) serán los mismos que en la Práctica 1.b (véase la Sección 3 de dicho guión de prácticas).

3. Componentes de los Algoritmos

Los algoritmos de esta práctica tienen en común las siguientes componentes:

- *Esquema de representación*: Se seguirá la representación binaria basada en un vector s de tamaño n con valores en $\{0,1\}$ que indica la selección o la eliminación de las características, explicada en las transparencias del seminario.
- *Función objetivo*: Será el porcentaje de acierto del clasificador 3-NN generado a partir de la selección de características codificada en la solución actual. El objetivo será maximizar esta función.
- *Algoritmo de búsqueda local*: Se considerará la búsqueda local (BL) que sigue el enfoque del primer mejor vecino propuesta en la Práctica 1.b. Se detendrá la ejecución del algoritmo bien cuando no se encuentre mejora en todo el entorno o bien cuando se hayan evaluado 15000 soluciones distintas.

A continuación veremos las particularidades de cada algoritmo.

3.1. Búsqueda Multiarranque Básica

Algoritmo

El algoritmo BMB consistirá simplemente en generar un determinado número de soluciones aleatorias iniciales y optimizar cada una de ellas con el algoritmo de BL considerado. Se devolverá la mejor solución encontrada en todo el proceso.

Valores de los parámetros y ejecuciones

En cada ejecución del algoritmo se generarán 25 soluciones iniciales aleatorias.

3.2. GRASP

Algoritmo

El algoritmo GRASP constará de dos componentes: construcción de soluciones *greedy* probabilísticas y optimización de las mismas mediante el algoritmo de BL. El algoritmo *greedy* probabilístico a considerar es el explicado en las transparencias del Seminario 2.b. Una vez generada cada solución *greedy* probabilística inicial se aplicará el algoritmo de BL sobre ella. Finalmente, se devolverá la mejor solución encontrada.

Valores de los parámetros y ejecuciones

En cada ejecución del algoritmo se generarán 25 soluciones iniciales aleatorias. El parámetro α que determina el umbral de tolerancia de calidad para la construcción de la lista restringida de candidatos tomará valor 0.3.

3.3. Búsqueda Local Reiterada (ILS)

Algoritmo

El algoritmo ILS consistirá en generar una solución inicial aleatoria y aplicar el algoritmo de BL sobre ella. Una vez obtenida la solución optimizada, se estudiará si es mejor que la mejor solución encontrada hasta el momento y se realizará una mutación sobre la mejor de estas dos, volviendo a aplicar el algoritmo de BL sobre esta solución mutada. Este proceso se repite un determinado número de veces, devolviéndose la mejor solución encontrada en todo el proceso. Por tanto, se sigue el *criterio del mejor* como criterio de aceptación de la ILS.

Tal y como se describe en las transparencias del Seminario 2.b, el operador de mutación de ILS estará basado en un operador de vecino que provoque un cambio más brusco en la solución actual que el considerado en la BL. Para ello, usaremos un operador que cambia conjuntamente la selección de t características *escogidas aleatoriamente* en la solución. El valor de s controlará el grado de cambio realizado sobre una solución.

Valores de los parámetros y ejecuciones

En cada ejecución del algoritmo se aplicará 25 veces el algoritmo de BL, la primera vez sobre una solución inicial aleatoria y las 24 restantes sobre soluciones mutadas. Se usará un valor $t=0.1 \cdot n$ en el operador de mutación, es decir, se cambiará el valor de selección de un 10% de las características.

4. Tablas de Resultados a Obtener

Se diseñará una tabla para cada algoritmo (3-NN, SFS, BMB, GRASP e ILS) donde se recojan los resultados de la ejecución de dicho algoritmo al conjunto de casos del problema. Tendrá la misma estructura que la Tabla 5.1 del guión de la Práctica 1.b.

Finalmente, se construirá una tabla de resultados global que recoja los resultados medios de calidad y tiempo para todos los algoritmos considerados, tal como se muestra en la tabla 4.1. Aunque en la tabla que sirve de ejemplo se han incluido todos los algoritmos considerados en esta práctica, naturalmente sólo se incluirán los que se hayan desarrollado.

Tabla 4.1: Resultados globales en el problema de la SC

	Wdbc			Movment Libras			Arrhythmia		
	% clas	% red	T	% clas	% red	T	% clas	% red	T
3-NN	x	0	x	x	0	x	x	0	x
SFS	x	x	x	x	x	x	x	x	x
BMB	x	x	x	x	x	x	x	x	x
GRASP	x	x	x	x	x	x	x	x	x
ILS	x	x	x	x	x	x	x	x	x

A partir de los datos mostrados en estas tablas, el alumno realizará un análisis de los resultados obtenidos, que influirá significativamente en la calificación de la práctica. En dicho análisis se deben comparar los distintos algoritmos en términos de las tasas de clasificación y de reducción obtenidas (capacidad del algoritmo para obtener soluciones de calidad) y del tiempo requerido para obtener las soluciones (rapidez del algoritmo). Se comparará el rendimiento de las metaheurísticas entre sí, así como con respecto a los algoritmos de referencia, el 3-NN original y el SFS.

5. Documentación y Ficheros a Entregar

Además de la documentación detallada en la Sección 6 del guión de la Práctica 1.b, en lo referente al punto 4 se incluirá, al menos, la siguiente información:

- Esquema de representación de soluciones empleado.
- Descripción en pseudocódigo de la función objetivo.
- Descripción en pseudocódigo del proceso de generación de soluciones aleatorias (usado en la BMB y en la ILS).
- Descripción en pseudocódigo del algoritmo de BL empleado, incluyendo el método de exploración del entorno, el operador de generación de vecino y su factorización.

En lo que respecta al punto 5, se incluirá la siguiente información:

- Descripción en pseudocódigo del esquema de búsqueda seguido por cada algoritmo (BMB, GRASP e ILS).
- Además se detallarán, al menos, las siguientes componentes particulares de cada algoritmo:
 - Para el algoritmo GRASP, descripción en pseudocódigo del mecanismo de generación de soluciones *greedy* probabilísticas, indicando el modo de obtención de la lista restringida de candidatos.
 - Para el algoritmo ILS, descripción en pseudocódigo del operador de mutación empleado.

En el punto 6, se incluirá una breve descripción del algoritmo de comparación, el *Greedy*.

Como recomendación, el apartado 4 debería describirse en un máximo de dos páginas. En el apartado 5, el número total de páginas para describir cada algoritmo (incluyendo el pseudocódigo del esquema de búsqueda y de las componentes particulares) sería de una página para BMB e ILS, y de dos páginas para GRASP.

Se recuerda que **la documentación nunca deberá incluir listado total o parcial del código fuente en caso de haberlo implementado.**

En lo referente al **desarrollo de la práctica**, se seguirán los mismos criterios descritos en la Sección 6 del guión de la Práctica 1.b. El **método de evaluación** será el descrito en la Sección 7 de dicho guión.