

**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 5.b:**  
**Búsquedas Híbridas para el Problema**  
**de la Selección de Características**

Curso 2015-2016

Tercer Curso del Grado en Ingeniería Informática

## Práctica 5.b

# Búsquedas Híbridas para el Problema de la Selección de Características

## 1. Objetivos

El objetivo de esta práctica es estudiar el funcionamiento de los Algoritmos Meméticos (AMs). Para ello, se requerirá que el alumno adapte estos métodos para resolver el problema de la selección de características 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**. La fecha límite de entrega será el **Lunes 11 de Julio 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/método de mejora*: 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 **cuando se haga una única iteración, se encuentre o no mejora en el entorno.**

#### Algoritmos

El AM consistirá en hibridar el algoritmo genético generacional (AGG) de la Práctica 3.b con la BL de la Práctica 1.a. Se estudiarán las tres posibilidades de hibridación siguientes:

1. AM-(10,1.0): Cada 10 generaciones, se aplica la BL sobre todos los cromosomas de la población.
2. AM-(10,0.1): Cada 10 generaciones, se aplica la BL sobre un subconjunto de cromosomas de la población seleccionado aleatoriamente con probabilidad  $p_{LS}$  igual a 0.1 para cada cromosoma.
3. AM-(10,0.1mej): Cada 10 generaciones, aplicar la BL sobre los  $0.1 \cdot N$  mejores cromosomas de la población actual ( $N$  es el tamaño de ésta).

#### Valores de los parámetros y ejecuciones

El tamaño de la población del AGG será de 10 cromosomas. Las probabilidades de cruce y mutación serán 0,7 y 0,001 (por gen) en ambos casos. El criterio de parada consistirá en realizar **15000 evaluaciones de la función objetivo**, incluidas por supuesto las de la BL.

### 4. Tablas de Resultados a Obtener

Se diseñará una tabla para cada algoritmo (3-NN, SFS, AM-(10,1.0), AM-(10,0.1) y AM-(10,0.1mej)) 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			Movement_Libras			Arrhythmia		
	% clas	% red	T	% clas	% red	T	% clas	% red	T
<b>3-NN</b>	x	0	x	x	0	x	x	0	x
<b>SFS</b>	x	x	x	x	x	x	x	x	x
<b>AM-(10,1.0)</b>	x	x	x	x	x	x	x	x	x
<b>AM-(10,0.1)</b>	x	x	x	x	x	x	x	x	x
<b>AM-(10,0.1mej)</b>	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 del mejor resultado individual obtenido (capacidad del algoritmo para obtener soluciones de calidad), los resultados medios (robustez del algoritmo) y el 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, tanto a nivel de los casos individuales como desde una perspectiva global.

## 5. Documentación y Ficheros a Entregar

Además de la documentación detallada en la Sección 6 del guión de la Práctica 1.a, 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 algoritmo de BL empleado, incluyendo el método de exploración del entorno y el operador de generación de vecino.

En el punto 5, se incluirá la descripción en pseudocódigo del esquema de búsqueda seguido por cada algoritmo (AM-(10,1.0), AM-(10,0.1) y AM-(10,0.1mej)).

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.

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.a. El **método de evaluación** será el descrito en la Sección 7 de dicho guión.