

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 4.b:
Optimización basada en Colonias de Hormigas para el
Problema de la Selección de Características

Curso 2015-2016

Tercer Curso del Grado en Ingeniería Informática

Práctica 4.b

Optimización basada en Colonias de Hormigas para el Problema de la Selección de Características

1. Objetivos

El objetivo de esta práctica es estudiar el funcionamiento de los algoritmos de Optimización basada en Colonias de Hormigas (OCH). Para ello, se requerirá que el alumno adapte dos variantes de estos métodos, el Sistema de Colonias de Hormigas y el Sistema de Hormigas Max-Min, ambos incorporando búsqueda local (SCH-BL y SHMM-BL), para resolver el problema de la selección de características descrito en las transparencias del Seminario 2.a. Los resultados obtenidos deberán compararse 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 **2 puntos**, distribuidos de la siguiente forma: SCH-BL (1 punto) y SHMM-BL (1 punto).

La fecha límite de entrega será el **Jueves 9 de Junio 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.

En este caso, el código proporcionado se denomina *ACOTSP.V1.03.tgz* y se ha obtenido de la web de la metaheurística OCH (<http://www.aco-metaheuristic.org/>). Implementa los principales algoritmos OCH existentes para la resolución del problema

del Viajante de Comercio (TSP). En el fichero README incluido en el paquete ACOTSP.V1.03 puede encontrarse una descripción del contenido e indicaciones sobre cómo compilar y ejecutar las distintas versiones de los algoritmos.

El código ha sido desarrollado en ANSI C bajo Linux. Una vez compilado usando el correspondiente *make* se genera un ejecutable llamado *acotsp* que permite recibir una serie de parámetros desde la línea de comandos. Puesto que ha sido implementado en ANSI C no es difícil de compilar en Windows teniendo en cuenta la información contenida en el fichero *Makefile*.

Los parámetros más importantes para las implementaciones del SCH y el SHMM son los siguientes:

- *-tries*: número de ejecuciones a realizar.
- *-ants*: número de hormigas en la colonia.
- *-time*: tiempo de ejecución (en segundos) empleado como criterio de parada.
- *-alpha* y *-beta*: parámetros de la regla de transición.
- *-rho*: parámetro de evaporación de los rastros de feromona.
- *-localsearch* y *-nnls*: búsqueda local y número de vecinos a generar, respectivamente.
- *-q0*: parámetro q_0 de la regla de transición del SCH.

Así, por ejemplo, una llamada correcta al programa SCH-BL para la instancia *eil76.tsp* del TSP cuyo óptimo es 538 sería:

```
./acotsp -tsplibfile eil76.tsp -tries 5 -acs -localsearch 3 -nnls 40 -alpha 1 -beta 2  
-rho 0.2 -q0 0.8 -time 300 -optimum 538 -ants 15
```

y en el caso del SHMM-BL:

```
./acotsp -tsplibfile eil76.tsp -tries 5 -mmas -localsearch 3 -nnls 40 -alpha 1 -beta 2  
-rho 0.2 -time 300 -optimum 538 -ants 15
```

Los métodos desarrollados serán ejecutados sobre una serie de casos del problema de la selección de características. 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

Algoritmos

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

- *Esquema de representación*: Las componentes de la solución L son las n características del problema. El grafo de construcción es un grafo totalmente conexo que incluye n nodos (f_1, \dots, f_n) y es totalmente conexo.
- *Restricciones*: No se pueden formar ciclos, es decir, no se puede escoger una característica más de una vez.
- *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.
- *Rastros de feromona*: Tendremos dos tipos distintos. Se asocia un rastro de feromona $\tau\text{-car}_{f_i}$ a cada nodo del grafo f_i . Indica la preferencia memorística de seleccionar la característica correspondiente. Se asocia otro rastro $\tau\text{-num-car}_{nc}$ a cada valor posible del número de características $nc \in \{1, \dots, n\}$ a seleccionar por cada hormiga.
- *Información heurística*: Se define como la preferencia individual de seleccionar la característica f_i , calculada como la cantidad de información de la variable f_i con respecto a las clases $\eta_{fi}=I(C, f_i)$. Para una explicación más detallada, consultar las transparencias del Seminario 4.
- *Proceso constructivo*: Antes de iniciar el proceso constructivo, se genera aleatoriamente el número de características nc que la hormiga concreta va a seleccionar en esta iteración usando la distribución de probabilidad de los rastros de feromona $\tau\text{-num-car}$.
En cada uno de los $n-1$ pasos necesarios para construir la solución ($k=1$ hasta nc), se repiten los tres pasos siguientes:
 - Se construye la lista de candidatos factibles con los nodos asociados a las variables no seleccionadas: $J(k) = \{f_i \in \{1, \dots, n\} \mid f_i \notin L\}$.
 - Se escoge un nodo f_j de la lista con la regla de transición (en función de la feromona $\tau\text{-car}_{f_i}$ y la heurística η_{fi}). Se almacena en L , $L[k] \leftarrow f_j$.
- *Reinicialización*: No es necesario emplear reinicialización.
- *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 aplica sobre las soluciones generadas por todas las hormigas en cada iteración del algoritmo OCH, antes de actualizar la feromona.** Se detendrá la ejecución del algoritmo después de haber realizado **una única iteración**, se haya encontrado mejora en el entorno o no.

Valores de los parámetros y ejecuciones

Los valores considerados para los parámetros de los algoritmos serán los siguientes. Se considerarán $m=10$ hormigas en la colonia. El criterio de parada consistirá en realizar **15000 evaluaciones** de la función objetivo, incluidas por supuesto la de la BL. Los parámetros de la regla de transición serán $\alpha=1$ y $\beta=2$.

El valor inicial de los rastros de feromona $\tau\text{-car}_{fi}$ será $\tau\text{-car}_{fi 0}=10^{-6}$ para el SCH-BL y $\tau\text{-car}_{fi 0}=\tau_{\max}$ para el SHMM-BL. En cambio, los rastros de feromona $\tau\text{-num-car}_{nc}$ se inicializarán a $1/N_c$ (N_c = número de características del problema). El parámetro de evaporación de feromona valdrá $\rho=0.2$ en los dos algoritmos de OCH. El aporte de feromona a realizar por la hormiga que generó solución S será $\Delta\tau^S=C(S)$ en los dos algoritmos, donde $C(S)$ es el valor de la función objetivo, es decir, el porcentaje de acierto del clasificador 3-NN generado a partir de las características seleccionadas en S .

Los límites de los rastros de feromona del SHMM-BL se calcularán y actualizarán de la siguiente forma. Antes de comenzar la ejecución del algoritmo se generará y evaluará una solución inicial aleatoria, inicializando τ_{\max} al valor $C(S_{\text{aleatoria}})/\rho$ y τ_{\min} a $\tau_{\max}/500$. Luego se inicializarán los rastros de feromona a τ_{\max} y comienza la ejecución del algoritmo. Posteriormente, cada vez que se actualice la mejor solución global en una iteración del algoritmo de OCH se actualizarán también los valores de τ_{\max} a $C(S_{\text{mejor_global}})/\rho$ y de τ_{\min} a $\tau_{\max}/500$. Justo después de actualizar los límites se recorrerá la matriz de rastros de feromona para truncar los valores que excedan τ_{\max} a τ_{\max} y los que queden por debajo de τ_{\min} a τ_{\min} .

Por último, en el caso del SCH-BL, el valor del parámetro q_0 en la regla de transición será 0.8 y el del parámetro de evaporación ϕ en la regla de actualización local de feromona será 0.2.

4. Tablas de Resultados a Obtener

Se diseñará una tabla para cada algoritmo (SCH-BL y SHMM-BL) 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
3-NN	x	0	x	x	0	x	x	0	x
SFS	x	x	x	x	x	x	x	x	x
AGG	x	x	x	x	x	x	x	x	x
AGE	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 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.b, en lo referente al punto 4 se incluirá, al menos, la siguiente información:

- a) Esquema de representación de soluciones empleado.
- b) Descripción en pseudocódigo de la función objetivo.
- c) Definición de los rastros de feromona y descripción en pseudocódigo del cálculo de la información heurística.
- d) Descripción en pseudocódigo del proceso constructivo para generar soluciones.
- e) Descripción en pseudocódigo del algoritmo de BL empleado, incluyendo el método de creación de la lista de candidatos, el 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:

- a) Descripción en pseudocódigo del esquema de búsqueda seguido por cada algoritmo (SCH-BL y SHMM-BL).
- b) Mecanismos de actualización y evaporación de feromona.

En el punto 6, se incluirá una breve descripción del algoritmo de comparación, el *Greedy SFS*. Como recomendación, el apartado 4 debería describirse en un máximo de tres páginas y los apartados 5 y 6 en un máximo de 2 páginas cada uno.

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.