

Documentación de la práctica de búsqueda local

Laboratorio de Inteligencia Artificial

2º Cuatrimestre - curso 2010/2011



Enginyeria en Informàtica

Departament de Llenguatges i Sistemes Informàtics



FIB

Facultat d'Informàtica
de Barcelona

UNIVERSITAT POLITÈCNICA DE CATALUNYA

Índice general

1. Organización, evaluación y entrega	2
2. Objetivos de aprendizaje	3
3. El problema	4
3.1. Elementos del problema	4
3.2. El problema	5
3.3. Criterios de la solución	5
3.4. Tareas a realizar	5
3.5. Experimentos	6
4. Guión de la práctica	8
5. Planificación de tareas	12
6. Experimentación	14
6.1. ¿Que es un experimento?	14
6.2. Diseño experimental	14
6.3. Pasos para el diseño de un experimento	14
6.4. Ejemplos de experimentación	15
6.4.1. TSP: Influencia de la solución inicial	15
6.4.2. TSP: Influencia del tamaño del problema	18
7. Rúbrica de evaluación	20

Organización, evaluación y entrega

Esta es la documentación de la práctica de búsqueda local, en este documento tenéis:

- Los objetivos de aprendizaje de la práctica correspondientes al temario de la asignatura
- La descripción del problema que debéis resolver, los elementos que debéis implementar en la solución del problema y los experimentos que deberéis realizar con vuestra solución
- Lo que tenéis que incluir en el informe que deberéis entregar como resultado de la práctica
- La planificación semanal de la práctica incluyendo los objetivos que debéis ir cubriendo cada semana y la dedicación en horas estimada.
- Documentación sobre cómo realizar experimentos
- Rúbrica de evaluación de la práctica

La práctica se debe hacer **preferentemente en parejas**. Intentad no hacerla solos ya que os llevará mucho más trabajo.

La práctica se debe desarrollar en **java** utilizando las clases que tenéis disponibles en la página web de laboratorio.

Planificad bien el desarrollo de la práctica y no lo dejéis todo para el último día, ya que no seréis capaces de acabarla y hacer un buen trabajo. En este documento tenéis indicaciones sobre el desarrollo de la práctica que os ayudará a planificar el trabajo.

En la evaluación de la práctica valoraremos sobre todo la calidad del análisis de cada experimento y los comentarios y conclusiones. Leed el capítulo de la rúbrica de evaluación para una descripción detallada de los criterios de evaluación. La garantía de una buena nota es seguir los criterios que aparecen en la rúbrica.

La entrega del informe y los fuentes de vuestro programa se hará el día **4 de abril** en formato electrónico según las instrucciones que aparecerán en el racó.

Objetivos de aprendizaje

El objetivo de esta práctica es resolver un problema mediante algoritmos de **búsqueda local**.

En esta práctica se ha de demostrar ser capaz de razonar sobre la naturaleza del problema y plantearlo como un problema de búsqueda local, solucionándolo con la librería de algoritmos que tenéis a vuestra disposición.

Los objetivos específicos que se tienen que cubrir con el desarrollo de la práctica son los siguientes:

- Razonar sobre qué elementos son los necesarios para plantear el espacio de soluciones del problema y como deberían representarse para solucionarlo mediante un algoritmo de búsqueda local. Se han de tener en cuenta las necesidades espacio/temporales de los algoritmos para elegir la representación.
- Determinar qué formas existen para generar una solución inicial del problema y razonar y justificar cuales son las más adecuadas. Se han de elegir varias alternativas para poder experimentar y probar las decisiones empíricamente.
- Determinar qué operadores de transformación sobre la solución son posibles en el problema y qué combinación de estos operadores es adecuada para solucionar el problema considerando: El factor de ramificación, la posibilidad de poder generar cualquier solución.
- Determinar y justificar las funciones heurísticas que permitan optimizar la solución según unos criterios específicos. Comparar las funciones heurísticas empíricamente y razonar sobre los resultados.
- Plantear un conjunto de experimentos que muestren la ejecución de los algoritmos de búsqueda Hill Climbing y Simulated Annealing en escenarios diferentes. Escoger los experimentos que demuestren el funcionamiento de los elementos que han elegido (solución inicial, operadores, función heurística) en diferentes situaciones del problema. Debéis de ser capaces de justificar su elección y comparar lo que se esperaba intuitivamente con el experimento con los resultados reales.
- Experimentar con los parámetros del Simulated Annealing para escoger los más adecuados en los experimentos.
- Comparar el comportamiento del algoritmo de Hill Climbing y el de Simulated Annealing respecto a la calidad de sus soluciones.

Respecto a la **evaluación**, tenéis disponible una rúbrica que indica los criterios que se usarán para valorar la práctica y una descripción de cada uno de los niveles de valoración para cada criterio.

El problema

Un problema de logística al que se enfrentan muchas empresas es la manera de hacer llegar mercancías a los centros de fabricación/venta/distribución desde los almacenes de materiales/productos. Habitualmente hay unos pocos almacenes y muchos centros de fabricación/venta/distribución. Estos suelen hacer peticiones al almacén indicando qué productos y cantidades son necesarias y cual es el plazo límite de entrega. Este plazo no tiene por que ser estricto, pero no cumplirlo puede suponer una pérdida o un aumento de costes.

3.1. Elementos del problema

Supondremos que una compañía tiene que gestionar las entregas desde un *almacén central* a un conjunto de *centros de producción*. Nosotros somos una compañía de transporte y nos subcontratan para hacer las entregas. El beneficio lo obtenemos por lo que nos pagan por hacer el transporte.

Estos centros realizan un conjunto de peticiones de productos al almacén al final del día. Cada petición indica el producto que se demanda, la cantidad (en kilogramos) y la hora límite de entrega. Supondremos que en nuestro problema hay 6 centros de producción. La cantidad de las peticiones asumiremos que es un número entero de 100 a 500 kg. en múltiplos de 100.

El almacén tendrá preparadas las peticiones que se han de repartir en el día y nosotros tenemos que organizar el transporte. Para ello disponemos de una flota de camiones. Cada camión tiene una capacidad de carga. Supondremos que tenemos camiones capaces de transportar 500, 1000 y 2000 kg.

Nosotros organizamos un horario para las entregas que supondremos que van desde las 8 de la mañana hasta las 5 de la tarde. La organización será de la siguiente manera:

- Para cada hora de entrega (de 8 a 17) y cada centro de producción (1 a 6) tendremos un camión.
- Este camión saldrá del almacén y llegará al centro de producción a la hora establecida.

Es decir, tenemos 60 transportes que programar. Asumiremos que tenemos camiones suficientes para realizar todos los transportes.

Para cada transporte deberemos decidir cual es la capacidad del camión que lo hace y qué peticiones transporta. Respecto a esas capacidades supondremos que tenemos un número específico de camiones de cada capacidad (en total sumarán los 60 camiones). De esta manera podremos experimentar la influencia del tipo de flota de camiones que tenemos en el problema.

Como hemos comentado, el beneficio lo obtendremos por el pago del transporte. La tabla de precios del envío de cada petición es la siguiente:

Peso	Precio
100 y 200 kg	<i>peso</i> euros
300 y 400 kg	$1,5 \times \textit{peso}$ euros
500 kg	$2 \times \textit{peso}$ euros

En el caso de que una petición llegue con retraso con respecto a la hora indicada, el precio que cobraremos se reducirá en un 20 % por cada hora de retraso (es decir, si el paquete se retrasa más de cinco horas nos tocará a nosotros pagar siguiendo la misma proporción). Evidentemente el paquete puede llegar antes de hora, pero en este caso no nos pagan de más. Si una petición no se puede entregar en el día, el coste para nosotros será el 20 % por cada hora de retraso hasta las 5 de la tarde, más el precio de entrega de la petición.

3.2. El problema

Dada una lista de peticiones de los centros de producción, donde cada una indica el centro de entrega, la cantidad y la hora límite de entrega y dado que vamos a realizar 60 transportes (uno por cada combinación de horas y centros de producción):

- Asignar el número máximo de peticiones a los transportes que vamos a realizar
- Determinar la capacidad del camión que va a realizar cada transporte

3.3. Criterios de la solución

Para obtener y evaluar la solución usaremos los siguientes criterios y restricciones:

- Tenemos una cantidad establecida de camiones de cada capacidad que suma el número total de transportes a realizar
- Tendremos dos criterios para evaluar la calidad de una solución:
 1. Maximizar la ganancia que obtenemos con los transportes
 2. Minimizar el valor absoluto de la diferencia entre la hora límite de entrega de la petición y la hora efectiva de entrega. Asumiremos que este valor para las peticiones no entregadas es la diferencia entre la hora límite de entrega y las 8 de la mañana del día siguiente.

3.4. Tareas a realizar

El desarrollo de la práctica implica realizar las siguientes tareas:

1. Implementar el problema de tal manera que se puedan generar problemas aleatorios. En este caso los elementos que varían son:
 - El número total de camiones de cada capacidad (siempre habrá 60 camiones en total).
 - El número de peticiones
 - Los destinos, pesos y horas de llegada de las peticiones. Para generar los valores de los pesos y las horas supondremos que siguen distribuciones de probabilidad uniformes y que se puede indicar como parámetro las probabilidades para cada valor de cada variable. La probabilidad de los destinos la consideraremos equiprobable.
2. Definir e implementar la representación del estado del problema para poder ser resuelto utilizando las clases del AIMA. Pensad bien en la representación, ha de ser eficiente en espacio y en tiempo.
3. Definir e implementar dos estrategias para generar la solución inicial.
4. Definir e implementar la función generadora de estados sucesores. Esto implica decidir el conjunto de operadores para explorar el espacio de búsqueda. Deberéis pensar y evaluar diferentes alternativas de conjuntos de operadores y justificar la elección de uno de ellos para realizar los experimentos. Deberéis implementar la función generadora de manera diferente para Hill Climbing y Simulated Annealing para que se puedan comparar sus tiempos de ejecución tal como se explicó en clase de laboratorio.

5. Definir e implementar dos funciones heurísticas, cada una de ellas implementará uno de los criterios de calidad de la solución indicados en el apartado anterior.

3.5. Experimentos

Deberéis hacer los siguientes experimentos:

1. Determinar qué conjunto de operadores da mejores resultados para una función heurística que optimice el primer criterio de calidad (3.3) con un escenario en el que el número de peticiones es 250 y la distribución de capacidades de los camiones, pesos de las peticiones y horarios de entrega es equiprobable. Deberéis usar el algoritmo de Hill Climbing. A partir de estos resultados deberéis fijar los operadores para el resto de experimentos. Escoged una estrategia de inicialización de entre las que proponéis.

Pensad que para explorar todas las posibles soluciones seguramente necesitaréis operadores que no influyen en los criterios de calidad de la solución definidos, por lo que tendréis que incluir más elementos en las funciones heurísticas para que estos operadores tengan efecto en la búsqueda.

2. Determinar qué estrategia de generación de la solución inicial da mejores resultados para la función heurística usada en el apartado anterior, con el escenario del apartado anterior y usando el algoritmo de Hill Climbing. A partir de estos resultados deberéis fijar también la estrategia de generación de la solución inicial para el resto de experimentos.
3. Determinar los parámetros que dan mejor resultado para el Simulated Annealing con el mismo escenario, usando la misma función heurística y los operadores y la estrategia de generación de la solución inicial escogidos en los experimentos anteriores.
4. Dado el escenario de los apartados anteriores, estudiad como evoluciona el tiempo de ejecución para hallar la solución en función del número de peticiones que tenemos. Para ello empezad con 200 peticiones e id aumentándolas de 50 en 50 hasta que veáis la tendencia. Usad el algoritmo de Hill Climbing y la misma heurística.
5. Dado el escenario del primer apartado con 200, 250 y 300 peticiones, estimad la diferencia entre la ganancia obtenida y el tiempo de ejecución para hallar la solución, usando dos heurísticas, la que habéis usado en los apartados anteriores y otra que optimice el segundo criterio de calidad propuesto (3.3) (notad que este segundo criterio no optimiza directamente la ganancia). Usad el algoritmo de Hill Climbing.
6. Dados los mismos escenarios que en el apartado anterior, estimad la diferencia entre la ganancia obtenida con el Hill Climbing y el Simulated Annealing para las dos heurísticas usadas en el apartado anterior y el tiempo de ejecución para hallar la solución de ambos algoritmos.
7. Para solucionar el problema podemos disponer de diferentes flotas de camiones. Dado el escenario del primer experimento, probaremos tres variaciones en las que tengamos un 50 % de camiones de un tamaño y un 25 % de cada uno de los otros dos. Usa el primer heurístico y el algoritmo de Hill Climbing. Intenta explicar la diferencia entre las soluciones de los distintos escenarios.
8. En el problema podemos tener algunas horas que tengan más peticiones que otras. Dado el escenario del primer experimento, genera variaciones en las que 4 de las horas tengan el doble de probabilidad de tener peticiones que el resto y compara las soluciones con las obtenidas en el primer experimento. Usa el primer heurístico y el algoritmo de Hill Climbing.

Para cada experimento deberéis hacer como mínimo 10 repeticiones y calcular valores medios. Podéis hacer gráficas y estadísticas que ilustren vuestras conclusiones, obviamente los que lo hagáis tendréis una mejor valoración.

Explicad y justificad todas las decisiones que toméis. Sacad conclusiones de los experimentos que incluyan **comentarios** sobre los resultados obtenidos en cada experimento. Explicad qué esperabais y qué habéis obtenido.

Guión de la práctica

Primera semana: Las clases del AIMA (21 de febrero)

Esta semana veréis en clase de laboratorio cómo utilizar las clases del AIMA para resolver problemas de búsqueda. Estas clases implementan la mayoría de los algoritmos que se han visto en clase y tenéis diferentes ejemplos de su utilización.

Tenéis ejemplos que resuelven problemas de búsqueda heurística y problemas de búsqueda local. Los primeros ejemplos que veréis son problemas de búsqueda heurística, en la práctica resolveréis un problema de búsqueda local, pero desde el punto de vista de realizar la implementación, los elementos que intervienen son los mismos, solo varía el significado que les asigna el algoritmo que se ha de ejecutar.

Fijaos que vosotros no tenéis que implementar los algoritmos, solo tenéis que identificar e implementar los elementos que estos necesitan. Estos elementos son los que se os han explicado en clase de teoría: Estado, operadores de búsqueda y función heurística.

Ejecutad los ejemplos que tenéis en las transparencias de laboratorio y echad un vistazo al código de los ejemplos que tenéis disponibles.

Segunda semana: Toma de contacto con el problema (28 de febrero)

A pesar de que todavía no podáis empezar a desarrollar la práctica deberíais haberos leído el enunciado y haberos hecho una idea sobre el problema.

Una cosa que deberíais hacer es reflexionar sobre el problema que describe el enunciado y preguntaros las mismas cosas que se preguntan en los problemas que hicisteis la primera clase de problemas.

- ¿Qué elementos intervienen en el problema?
- ¿Cuál es el espacio de búsqueda?
- ¿Qué tamaño tiene el espacio de búsqueda?
- ¿Qué es un estado inicial?
- ¿Qué condiciones cumple un estado final?
- ¿Qué operadores permiten modificar los estados?
- ¿Qué factor de ramificación tienen los operadores de cambio de estado?

A partir de las respuestas a esas preguntas os podéis plantear lo que necesitaréis para implementar la práctica con las clases del AIMA.

Un primer ejercicio consiste en que penséis qué estructura de datos debéis implementar para representar el estado. Es fundamental que penséis la representación teniendo en cuenta la mayor eficiencia espacial y temporal, ya que la búsqueda generará una gran cantidad de estados.

Podéis mirar el código de los ejemplos y prácticas anteriores que tenéis disponibles para haceros una mejor idea de lo que necesitareis implementar.

No dudéis en preguntar vuestras dudas a vuestros profesores de laboratorio o en el fórum de la práctica.

Tercera semana, implementación del estado (7 de marzo)

Esta semana ya deberíais tener claros los elementos que permiten definir el problema de la práctica. Ahora deberíais empezar a plantearos la implementación de la clase que representa el estado.

La implementación del estado incluye la decisión de qué estructura de datos es más adecuada para representar los elementos del problema. Debería ser eficiente ya que la exploración ha de generar un número bastante grande de estados.

Para ganar eficiencia espacial es una buena idea declarar estáticas las partes de la representación que no cambian, de manera que se compartan entre todas las instancias.

Deberéis implementar constructores que generen la solución inicial. El enunciado os pide que busquéis al menos dos estrategias para generarla.

Tenéis que pensar varias cosas sobre cómo generar la solución inicial, entre ellas cuál es el coste de generarla y cómo de buena es.

De cara a hacer los experimentos interesa ver como influye la bondad de la solución en el resultado de la búsqueda. Por ejemplo, podéis usar una estrategia que genera soluciones muy malas y al menos otra que presumiblemente sea mejor según algún criterio, deberéis ver si el número de pasos hasta llegar a la solución final y su calidad es diferente.

Dentro de la implementación del estado también esta incluida la implementación de los operadores de búsqueda. Tenéis que analizar qué conjunto de operadores es más conveniente. Observad que el factor de ramificación es importante, porque influye directamente en el tiempo para hallar la solución. Debéis pensar también que unos operadores que no generen suficientes alternativas pueden dar lugar a soluciones peores por no permitir explorar correctamente el espacio de búsqueda.

Pensad bien en las operaciones que se pueden hacer en el problema. A veces hacen falta varias para poder acceder a todo el espacio de búsqueda, también puede haber conjuntos alternativos de operadores. Deberéis sopesar el factor de ramificación y la conectividad que se obtiene entre las soluciones.

También deberíais ir escribiendo la documentación de la práctica a medida que vais haciendo cosas, no lo dejéis para el final.

Cuarta semana: Implementación de las clases para el AIMA (14 de marzo)

A estas horas ya deberíais tener implementado el estado del problema junto con sus operadores y las estrategias de inicio de la búsqueda.

Ahora necesitáis implementar el resto de clases que el AIMA usa para resolver el problema.

Implementar la clase que genera los estados sucesores es sencillo, solo tenéis que generar para un estado todos los estados accesibles posibles. Como ya tendréis implementados los operadores solo tendréis que decidir cómo se hace la generación de los nodos aplicando los operadores. El orden en que se generan es indiferente, solo tenéis que aseguraros de que para cada nodo se generan todos los sucesores accesibles. Recordad que la estrategia para generar los nodos ha de ser diferente para Hill Climbing que para Simulated Annealing. Para Hill Climbing tendréis que generar todas las posibles aplicaciones de los operadores al estado actual, mientras que para Simulated Annealing tendréis que escoger al azar un operador y generar solo un sucesor aplicando este operador con parámetros también al azar.

Con esto es suficiente, ya que es el algoritmo de búsqueda es el que se encargará de hacer la exploración y decidir qué nodos se expanden.

La clase que comprueba si se ha llegado a un estado final es la más sencilla de todas, en el caso de búsqueda local no es posible saber si se ha llegado al estado final, por lo que la función que implementa esta clase ha de retornar siempre **falso**. Podéis copiar esta clase de cualquiera de los ejemplos de búsqueda local que tenéis, todas son iguales.

La clase que debéis pensar más es la que calcula la función heurística.

Tenéis que pensar en qué miden las funciones heurísticas que describe el enunciado y decidir cómo se calculan a partir del estado.

Acordaos de que las funciones heurísticas que implementéis han de minimizarse. Pensad también que la diferencia entre maximizar y minimizar es un cambio de signo.

Quinta semana: Experimentos (21 de marzo)

Durante esta semana deberíais tener ya una implementación funcional de la práctica y comenzar a hacer experimentos.

Deberéis pensar en los diferentes escenarios que se pueden plantear con los elementos que tenéis. Tened en cuenta que el objetivo de los experimentos es obtener información que os permita responder a las preguntas que plantea el enunciado. Tened en cuenta que un mismo experimento os puede dar información para varias preguntas.

Seguid el orden de los experimentos que tenéis en el enunciado, éste os permitirá ir tomando decisiones sobre los diferentes elementos del problema e ir fijándolas para experimentos sucesivos.

Para ajustar los parámetros del Simulated Annealing escoged valores extremos y probad sus efectos. A partir de ellos podéis ir ajustando más sus valores explorando puntos intermedios hasta llegar a un valor que os parezca adecuado según el objetivo del problema. Tened en cuenta el significado de los parámetros para guiaros en vuestra exploración.

Haced suposiciones sobre cómo deberían ser las soluciones en cada experimento y comprobad los resultados que obtenéis. Comparad si los resultados corresponden con vuestras intuiciones e intentad justificar el resultado.

Tened en cuenta que para sacar conclusiones con fundamento deberéis ejecutar cada experimento varias veces. Para poder comparar deberéis ejecutar los algoritmos con datos iguales, para ello podéis generarlos fijando las semillas del generador de números aleatorios al crear los datos.

Podéis justificar la significatividad de los resultados que obtenéis utilizando los conocimientos que habéis adquirido en la asignatura de estadística. También podéis utilizar gráficos para ilustrar lo que sucede en los experimentos.

No os olvidéis de ir escribiendo la documentación a medida que hagáis los experimentos.

Sexta semana: La documentación final (28 de marzo)

En esta semana deberíais tener los resultados de los experimentos y escribir la documentación.

Algo que tenéis que tener presente es que la documentación ha de ser un reflejo de vuestro trabajo ya que es lo que servirá para que califiquemos vuestra práctica. Una mala documentación significa una mala nota.

La documentación deberá incluir:

- La descripción/justificación de la implementación del estado
- La descripción/justificación de los operadores que habéis elegido
- La descripción/justificación de las estrategias para hallar la solución inicial
- La descripción/justificación de las funciones heurísticas
- Para cada experimento:
 - Condiciones de cada experimento

- Resultados del experimento
- Qué esperabais y qué habéis obtenido
- Comparaciones
- Comentarios adicionales que os parezcan adecuados
- Comparación entre los resultados obtenidos con Hill Climbing y simulated annealing (no olvidéis explicar cómo habéis ajustado los parámetros para este último algoritmo).
- Respuestas razonadas a las preguntas del enunciado.

Planificación de tareas

Esta planificación os puede servir como orientación para organizar y repartir el trabajo de la práctica. Tened en cuenta que una planificación adecuada beneficiará la calidad de vuestra práctica.

Las horas asignadas al trabajo de la práctica según los créditos ECTS que le corresponde son 30 por cada componente del grupo, más las horas de las clases de laboratorio que son 6. En total la práctica corresponde a un esfuerzo de 72 horas durante seis semanas. Tened en cuenta que son horas ECTS, por lo que este es el tiempo de dedicación de un alumno medio para hacer la práctica a un nivel aceptable (digamos que una nota de notable) para llegar al sobresaliente hace falta algo más de dedicación.

Las tareas a desarrollar dentro de las semanas que indica el guión son las que aparecen en la siguiente tabla. También se indican las horas de dedicación aproximada del grupo en conjunto.

Semana 1 (2h lab + 6 horas trabajo externo)
<ol style="list-style-type: none"> Entender como funciona AIMA Mirar/ejecutar los ejemplos Leer y entender en enunciado
Semana 2 (2h lab + 12 horas trabajo externo)
<ol style="list-style-type: none"> Entender como funciona AIMA Pensar la representación y elementos del problema Diseñar la estructura de datos del estado Definir la interfaz de acceso a la estructura Documentar la estuctura de datos del estado
Semana 3 (2h lab + 12 horas trabajo externo)
<ol style="list-style-type: none"> Implementación del estado Implementación de la generación del estado inicial Implementación de los operadores de búsqueda Documentación
Semana 4 (2h lab + 10 horas trabajo externo)
<ol style="list-style-type: none"> Función generadora de sucesores Función que indica el estado final Cálculo de las funciones heurísticas Documentación

Semana 5 (2h lab + 10 horas trabajo externo)
<ol style="list-style-type: none"> 1. Planificación de los experimentos 2. Realizar los experimentos y recolectar resultados 3. Analizar los resultados de los experimentos
Semana 6 (2h lab + 10 horas trabajo externo)
<ol style="list-style-type: none"> 1. Documentación

Las horas de laboratorio de las dos primeras semanas os servirán para aprender como funciona el AIMA. El resto de horas de laboratorio las deberéis dedicar para las consultas al profesor de laboratorio, sincronización entre los componentes del grupo y desarrollo de la práctica.

La tarea de planificación de los experimentos en la quinta semana es muy importante. Una vez decididos los experimentos a realizar, estos se pueden ejecutar en batch y después analizar los resultados.

Es importante que os dividáis las tareas entre los dos miembros del grupo de manera adecuada. Muchas de las tareas se pueden hacer de manera paralela. Y recordad que dos personas delante de un PC durante una hora son una hora de trabajo, no dos.

Experimentación

6.1. ¿Que es un experimento?

“Un experimento es un *procedimiento* mediante el cual se trata de comprobar (confirmar, verificar o inventar) una o varias hipótesis relacionadas con un determinado fenómeno, mediante la manipulación de una o más variables que presumiblemente son su causa.”

(Wikipedia, <http://es.wikipedia.org/wiki/Experimento>)

6.2. Diseño experimental

“El diseño experimental es una técnica estadística que permite identificar y cuantificar las causas de un efecto dentro de un estudio experimental. En un diseño experimental se manipulan deliberadamente una o más variables, vinculadas a las causas, para medir el efecto que tienen en otra variable de interés. El diseño experimental prescribe una serie de pautas relativas a qué variables hay que manipular, de qué manera, cuántas veces hay que repetir el experimento y en qué orden para poder establecer con un grado de confianza predefinido la necesidad de una presunta relación de causa-efecto.”

(Wikipedia, http://es.wikipedia.org/wiki/Diseño_experimental)

6.3. Pasos para el diseño de un experimento

- Observación
- Planteamiento del problema de investigación
- Hipótesis: hipótesis nula (H_0) e hipótesis alterna
- Método (incluye la elección de los sujetos, para la conformación de la muestra; el procedimiento a seguir, es decir, el tratamiento a aplicar a los sujetos; las variables consideradas: variable dependiente, variable independiente, variables extrañas)
- Resultados: aquí se describen cuáles fueron las relaciones observadas entre las variables (si los valores de la variable independiente realmente influyeron significativamente sobre los de la variable dependiente, si hubo tantas variables extrañas como se pensaba o si surgieron otras), para lo cual se añaden a dicha descripción tanto gráficas (de barras, de pastel, etc.) como cuadros.

- Conclusiones

(Wikipedia, http://es.wikipedia.org/wiki/Diseño_experimental)

6.4. Ejemplos de experimentación

Para ilustrar como hacer una serie de experimentos dentro del tipo de problemas que se solucionan en esta práctica, vamos a hacer algunos experimentos con el problema del viajante de comercio usando la implementación que hay en las clases de AIMA.

6.4.1. TSP: Influencia de la solución inicial

Un elemento importante en un problema de búsqueda local es la influencia del punto de partida en el coste de la solución inicial. A priori no se puede determinar si un método de inicio es mejor o peor dado que se desconoce la forma que tendrá el espacio de búsqueda.

Para probar la hipótesis de que hay métodos de inicialización mejores que otros vamos a realizar un experimento probando diferentes métodos y evaluando sus resultados. Consideraremos tres métodos distintos de inicialización: Camino inicial ordenado según el identificador de la ciudad, camino inicial aleatorio y camino obtenido mediante una estrategia avariciosa.

Para poder hacer el experimento en condiciones debemos garantizar que todos los experimentos con cada hipótesis se realizan en las mismas condiciones. Para ello vamos a realizar diez réplicas del experimento donde probaremos cada método de inicialización. Para cada experimento individual utilizaremos la misma semilla de números aleatorios, de manera que cada réplica sea idéntica en las pruebas para cada método de inicialización.

En este caso podemos medir diferentes cosas en cada experimento, pero antes de hacerlo debemos pensar en qué es comparable y qué no.

Por un lado el valor de la solución de cada experimento no es comparable, ya que cada réplica partirá de condiciones diferentes. En este caso nos puede interesar para cada experimento cual de los métodos obtuvo la mejor solución.

También nos puede interesar el tiempo que tardaron en solucionarse los problemas y el número de pasos que se hicieron hasta llegar a la solución. Estas magnitudes son más comparables entre diferentes réplicas ya que deberían depender del tamaño del problema (en este caso las ciudades). Hay que fijarse también que estas dos medidas están altamente correlacionadas, por lo que nos llevarán a las mismas conclusiones.

Otra dificultad en nuestra experimentación es que uno de los métodos es aleatorio, por lo que diferentes ejecuciones del experimento con condiciones iguales dará diferentes resultados. Para poder comparar con los otros métodos que siempre darán la misma solución en las mismas condiciones podemos hacer la media de los resultados obtenidos de varias ejecuciones.

Podemos resumir las características de este experimento en la siguiente tabla:

Observación	Pueden haber métodos de inicialización que obtienen mejores soluciones
Planteamiento	Escogemos diferentes métodos de inicialización y observamos sus soluciones
Hipótesis	Todos los métodos de inicialización son iguales (H_0) o hay métodos mejores que otros
Método	<ul style="list-style-type: none">■ Elegiremos 10 semillas aleatorias, una para cada réplica■ Ejecutaremos 1 experimento para cada semilla para la inicialización ordenada y avariciosa■ Ejecutaremos 5 experimentos para cada semilla para la inicialización aleatoria y haremos medias de los resultados■ Experimentaremos con problemas de 40 ciudades■ Usaremos el algoritmo de Hill Climbing■ Mediremos diferentes parámetros para realizar la comparación

Una vez realizado el experimento debemos recolectar la información que queremos usar para probar

nuestras hipótesis. En las tablas 6.1, 6.2 y 6.3 podemos ver los resultados de un experimento como el que hemos definido.

A partir de estos datos podemos empezar a dar respuestas a nuestras hipótesis.

Por ejemplo, podemos comparar cada pareja de métodos de inicialización asumiendo que todos son igual de buenos (la hipótesis nula). Para ello podemos considerar que la probabilidad de que un método dé una solución mejor que otro para un experimento se distribuye como una función binomial y podemos comprobar cual es la probabilidad de que la hipótesis nula sea cierta (que entre dos métodos, los dos tienen la misma probabilidad de dar la mejor solución).

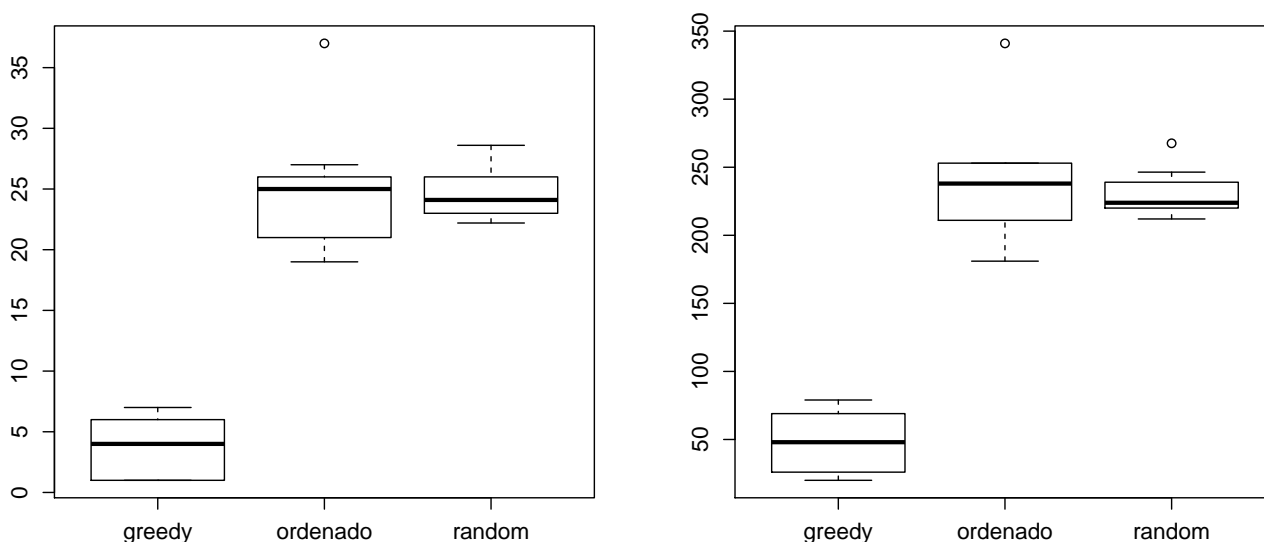
Para el caso de la inicialización avariciosa respecto a la ordenada tenemos que siempre la avariciosa da una solución mejor, por lo que deberíamos consultar en las tablas de la binomial cual es la probabilidad de que en 10 experimentos tengamos 10 experimentos donde avariciosa gane, asumiendo que son igual de buenos (tienen una probabilidad de 0,5 de dar el mejor resultado), con ello podemos comprobar que esa probabilidad es 0,001, o sea, bastante remota. Esto querría decir que no es cierto que sean iguales. Sucede lo mismo con la comparación entre el avaricioso y el aleatorio.

Para la comparación entre el ordenado y el aleatorio, tenemos que el ordenado es mejor 6 de las 10 veces. Si consultamos la binomial tendremos que la probabilidad es de 0,205, significativamente más alta que entre estas y la avariciosa.

Podemos comparar también las distribuciones estadísticas del número de pasos y el tiempo en hallar la solución, en la siguiente tabla tenemos la media y desviación de estas variables:

	Greedy	Orden	Aleatorio
Pasos	3.6 (2.41)	25 (4.98)	24.56 (2.96)
Tiempo	46.8 (21,52)	239.3 (42.16)	229.7 (16,8)

En la siguiente gráfica se puede apreciar mejor la distribución del número de pasos y el tiempo de ejecución:



Es obvio que el método avaricioso es mucho mas rápido que los otros dos. La explicación más plausible es que la solución inicial se encuentra muy cerca de la solución final si usamos este método. Los otros comienzan en una solución mas lejana y tienen que recorrer un camino más largo hasta un óptimo local.

Se nos puede ocurrir también comprobar si el número de pasos (o el tiempo) que necesitan los métodos ordenado y aleatorio son parecidos. Para ello podemos asumir que su distribución es normal y realizar un test de hipótesis para ver si la media de las distribuciones es idéntica. Esto se puede

Réplica	Greedy	Ordenado	Aleatorio
1	150	214	188
2	156	172	187,6
3	172	184	189,6
4	104	208	163,6
5	150	212	208,8
6	142	194	192,8
7	142	202	206,8
8	138	172	184
9	134	182	198,4
10	142	184	202

Figura 6.1: Valor de la solución

Réplica	Greedy	Ordenado	Aleatorio
1	1	21	26
2	1	26	22,2
3	7	26	26,6
4	1	19	24,4
5	6	23	24,2
6	5	27	23,8
7	1	24	22,8
8	6	37	28,6
9	4	21	23
10	4	26	24

Figura 6.2: Número de pasos

Réplica	Greedy	Ordenado	Aleatorio
1	28	253	239
2	25	243	212
3	79	239	246,4
4	20	181	227
5	71	211	223,6
6	51	253	221,2
7	26	225	216,2
8	69	341	267,6
9	54	210	220
10	45	237	224,2

Figura 6.3: Tiempo en hallar la solución

hacer con un test de t-Student par pares de muestras (lo podeis calcular con cualquier paquete estadístico). Si hacemos este test para el número de pasos usando una confianza del 0.95 obtenemos:

```
data: pasosOrdenado and pasosAleatorio
t = 0.3313, df = 9, p-value = 0.748
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-2.563969  3.443969
sample estimates:
mean of the differences:  0.44
```

Un valor de probabilidad de 0,748 nos dice que es bastante probable que la diferencia entre las medias de las muestras sea 0. Por lo tanto no hay diferencia significativa entre los pasos que necesitamos para llegar a un óptimo con ambas inicializaciones.

6.4.2. TSP: Influencia del tamaño del problema

También nos puede interesar tener una idea del aumento del coste de la búsqueda en función del tamaño del problema. Podemos experimentar con el TSP usando números de ciudades crecientes para hacernos una idea del aumento del tiempo de cálculo. Podemos asumir que el crecimiento del tiempo estará en función del factor de ramificación y del tamaño del espacio de búsqueda, en este caso en esta implementación del problema el factor de ramificación es cuadrático (intercambios de todos los pares de ciudad) y el tamaño del espacio de búsqueda es $O(nciudades!)$.

Aquí asumiremos que el tiempo para los problemas de un mismo tamaño es comparable, aunque dependerá mucho de la solución inicial. Haremos varias ejecuciones para cada tamaño y usaremos los valores medios para ajustar la función del tiempo:

Podemos resumir el experimento en:

Observación	El tiempo sigue una función creciente respecto al tamaño del problema
Planteamiento	Escogemos diferentes tamaños de problema y observamos sus tiempos de ejecución
Hipótesis	La función será al menos cuadrática respecto al tamaño del problema
Método	<ul style="list-style-type: none"> ■ Elegiremos 10 semillas aleatorias para cada tamaño del problema ■ Ejecutaremos 10 experimentos para cada tamaño y haremos medias de los resultados ■ Experimentaremos con problemas de 25, 50, 75 y 100 ciudades ■ Usaremos el algoritmo de Hill Climbing con inicialización avariciosa ■ Mediremos el tiempo de cómputo para realizar la comparación

Si hacemos el experimento podemos tener unos resultados distribuidos como los de la figura 6.4. En la siguiente tabla podemos ver la media y desviación del tiempo de ejecución (ms):

Tamaño	25	50	75	100
Media	13,8	98,2	245,7	1196,9
Desviación	5,07	33,34	89,49	807,66

En la figura 6.5 podemos ver el crecimiento del tiempo en función del tamaño. La línea marcada como 1 es el tiempo medio obtenido de los experimentos. La línea marcada como 2 sería el tiempo suponiendo que crece como una función cuadrática del tamaño del problema. La línea marcada como 3 sería el tiempo suponiendo que crece como una función cúbica del tamaño del problema.

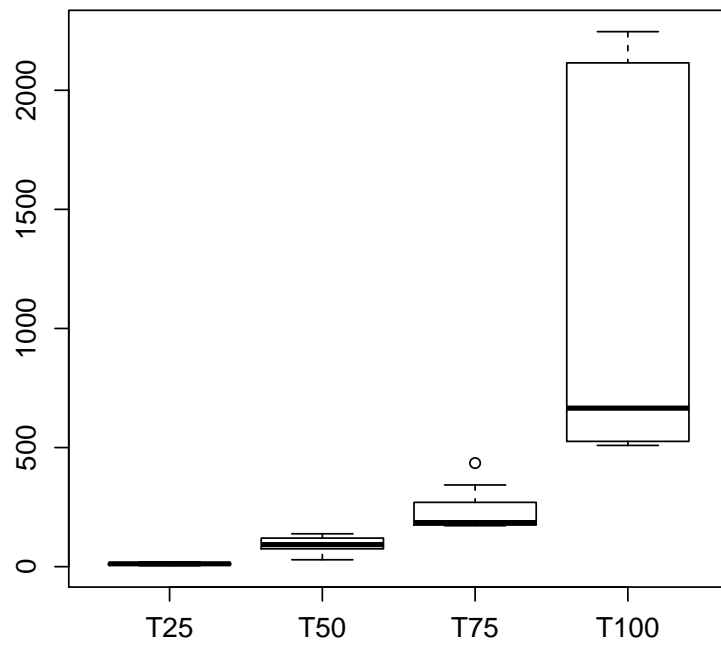


Figura 6.4: Distribucion del tiempo (ms) para cada tamaño del problema

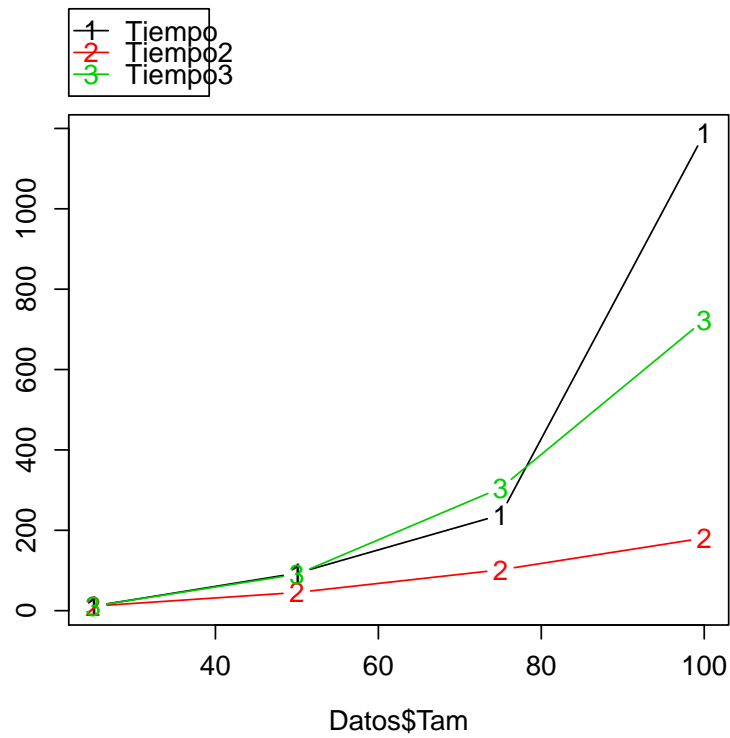


Figura 6.5: Función del tiempo (ms) respecto al tamaño del problema

Rúbrica de evaluación

Esta es la rúbrica de evaluación de la práctica. La corrección se hará según estos criterios y siguiendo las pautas que se detallan para cada nivel de evaluación.

Deberéis seguir estos criterios a la hora de escribir vuestra documentación y explicar qué habéis hecho en el desarrollo de la práctica y como lo habéis hecho.

Valoración			
Criterio	Mal	Regular	Bien
Parte descriptiva Identificación del problema	<ul style="list-style-type: none"> La descripción del problema es una copia de la descripción que hay en el enunciado. 	<ul style="list-style-type: none"> La descripción del problema es una copia de la descripción que hay en el enunciado. Análisis breve de las características del problema. 	<ul style="list-style-type: none"> La descripción del problema es algo más elaborada que la que ya hay en el enunciado. Análisis detallado de las características del problema.
Estado del problema y representación	<ul style="list-style-type: none"> No hay una descripción de los elementos del estado del problema. No hay un análisis del tamaño del espacio de búsqueda. Hay una descripción/justificación poco detallada de la representación del problema. La representación del problema es inadecuada y/o ineficiente. 	<ul style="list-style-type: none"> Hay una descripción breve de los elementos del estado del problema. Hay un análisis del tamaño del espacio de búsqueda. Hay una descripción/justificación poco detallada de la representación del problema. La representación del problema es adecuada y eficiente. 	<ul style="list-style-type: none"> Hay una descripción detallada los elementos del estado del problema. Hay un análisis del tamaño del espacio de búsqueda. Hay una descripción/justificación detallada de la representación del problema. La representación del problema es adecuada y eficiente.
Representación y análisis de los operadores	<ul style="list-style-type: none"> No hay una descripción de los operadores (sus condiciones de aplicabilidad y sus efectos). No hay un análisis del factor de ramificación de los operadores. No hay una explicación de la elección de los operadores. 	<ul style="list-style-type: none"> Hay una descripción breve de los operadores (sus condiciones de aplicabilidad y sus efectos). Hay un análisis del factor de ramificación de los operadores. No hay una explicación de la elección de los operadores. 	<ul style="list-style-type: none"> Hay una descripción detallada de los operadores (sus condiciones de aplicabilidad y sus efectos). Hay un análisis del factor de ramificación de los operadores. Hay una explicación de la elección de los operadores.

Valoración		Mal	Regular	Bien
Criterio	Análisis de la función heurística	<ul style="list-style-type: none"> No hay una explicación/análisis de los factores que intervienen en la heurística del problema. No hay una justificación de las funciones heurísticas escogidas. No hay una explicación de los efectos de las funciones heurísticas en la búsqueda. No hay una justificación de las ponderaciones que aparecen entre los elementos de las heurísticas. 	<ul style="list-style-type: none"> No hay una explicación/análisis de los factores que intervienen en la heurística del problema. Hay una justificación de las funciones heurísticas escogidas. Hay una explicación de los efectos de las funciones heurísticas en la búsqueda. No hay una justificación de las ponderaciones que aparecen entre los elementos de las heurísticas. 	<ul style="list-style-type: none"> Hay una explicación/análisis de los factores que intervienen en la heurística del problema. Hay una justificación de las funciones heurísticas escogidas. Hay una explicación de los efectos de las funciones heurísticas en la búsqueda. Hay una justificación de las ponderaciones que aparecen entre los elementos de las heurísticas.
	Elección y generación del estado inicial	<ul style="list-style-type: none"> No hay una descripción del algoritmo para calcular las soluciones iniciales. No hay una explicación/justificación de las elecciones de solución inicial (bondad de la solución, coste de hallar la solución). 		<ul style="list-style-type: none"> Hay una descripción del algoritmo para calcular las soluciones iniciales. Hay una explicación/justificación de las elecciones de solución inicial (bondad de la solución, coste de hallar la solución).

Valoración		Mal	Regular	Bien
Criterio	Parte experimental			
	Sobre los experimentos en general	<ul style="list-style-type: none"> ▪ No son informativos y/o solo se ha hecho una ejecución de cada experimento. ▪ No están encaminados a destacar las diferencias entre los elementos comparados. ▪ No se ha escogido un criterio uniforme para comparar los experimentos. ▪ No se explican, ni analizan los resultados (tablas, estadísticas, gráficos). ▪ No se compara lo que se esperaba con lo obtenido. 	<ul style="list-style-type: none"> ▪ Son informativos y se han hecho varias ejecuciones de cada experimento. ▪ Están encaminados a destacar las diferencias entre los elementos comparados. ▪ No se explican, ni analizan los resultados (tablas, estadísticas, gráficos) o el análisis no tiene mucho sentido. ▪ No se compara lo que se esperaba con lo obtenido. 	<ul style="list-style-type: none"> ▪ Son informativos y se han hecho varias ejecuciones de cada experimento. ▪ Están encaminados a destacar las diferencias entre los elementos comparados. ▪ Se explican y analizan correctamente los resultados (al menos con tablas, mejor con estadísticas y gráficos). ▪ Se compara lo que se esperaba con lo obtenido.
Influencia de la solución inicial		<ul style="list-style-type: none"> ▪ No se muestra claramente la diferencia entre las soluciones iniciales propuestas. 		<ul style="list-style-type: none"> ▪ Se muestra claramente la diferencia entre las soluciones iniciales propuestas: <ul style="list-style-type: none"> • En función del coste temporal de la búsqueda y el ahorro justificable por partir desde cada solución inicial. • En función de la bondad de las soluciones comparando si partir de mejores soluciones iniciales lleva a mejores soluciones finales.
Influencia de los operadores		<ul style="list-style-type: none"> ▪ No se muestra claramente la diferencia entre los conjuntos de operadores. 		<ul style="list-style-type: none"> ▪ Se muestra claramente la diferencia entre los conjuntos de operadores: <ul style="list-style-type: none"> • En función del coste temporal de la búsqueda. • En función de la bondad de las soluciones.

Valoración		Mal	Regular	Bien
Criterio	Influencia de la función heurística	<ul style="list-style-type: none"> ▪ No se muestra claramente la diferencia entre las funciones heurísticas escogidas. ▪ No se experimentan las ponderaciones de los criterios de las funciones heurísticas. 		<ul style="list-style-type: none"> ▪ Se muestra claramente la diferencia entre las funciones heurísticas: <ul style="list-style-type: none"> • En función del coste temporal de la búsqueda. • En función de la bondad de las soluciones. ▪ Se realizan experimentos para explorar la influencia de diferentes ponderaciones para los criterios de las funciones heurísticas.
	Comparación de algoritmos	<ul style="list-style-type: none"> ▪ No se hace un ajuste de los parámetros del Simulated Annealing. ▪ No se hace una comparación de los dos algoritmos de búsqueda en función de los experimentos. 	<ul style="list-style-type: none"> ▪ Se ajustan los parámetros del Simulated Annealing pero sin método alguno. ▪ Se hace una comparación pobre de los algoritmos de búsqueda (HC,SA) sin ningún criterio definido. 	<ul style="list-style-type: none"> ▪ Se ajustan los parámetros del Simulated Annealing explorando sistemáticamente los valores de sus parámetros. ▪ Se hace una comparación de los algoritmos de búsqueda en función de: <ul style="list-style-type: none"> • El coste temporal de la búsqueda. • La bondad de las soluciones.
	Escenarios propuestos y preguntas del enunciado	<ul style="list-style-type: none"> ▪ No se experimentan los escenarios planteados en el enunciado. ▪ No se responden a la preguntas planteadas en el enunciado. 	<ul style="list-style-type: none"> ▪ Se hacen los experimentos planteados en el enunciado pero las explicaciones sobre los resultados son breves y poco justificadas. ▪ Se responden a las preguntas del enunciado pero sin apoyar las explicaciones en los resultados de los experimentos. 	<ul style="list-style-type: none"> ▪ Se hacen los experimentos planteados en el enunciado, las explicaciones sobre los resultados son adecuadas y razonadas. ▪ Las explicaciones comparan/justifican los resultados que se esperaban con el escenario y los resultados obtenidos con los experimentos. ▪ Se responden a las preguntas del enunciado apoyando las explicaciones con los resultados de los experimentos y/o realizando experimentos adicionales.

Valoración		Mal	Regular	Bien
Criterio				
Calidad de la práctica				
Documentación	<ul style="list-style-type: none"> La práctica tiene una mala presentación. El documento esta mal estructurado y/o es ilegible. La documentación no demuestra una buena planificación del trabajo. La documentación no refleja el esfuerzo realizado. 			
	<ul style="list-style-type: none"> La práctica tiene una presentación adecuada. El documento está estructurado según los puntos que indica la práctica. La explicaciones están bien elaboradas y son inteligibles. La documentación demuestra una buena planificación del trabajo y un seguimiento del guión de la práctica. La documentación refleja el esfuerzo realizado. 			

Valoración		Mal	Regular	Bien
Criterio	Elaboración de informes de prácticas			
Ortografía y gramática	<ul style="list-style-type: none">El documento está plagado de errores de ortografía y gramaticales. Desde este punto de vista, es un documento impresentable.	<ul style="list-style-type: none">El documento tiene pocas faltas de ortografía y gramaticales.	<ul style="list-style-type: none">El documento no tiene faltas de ortografía, ni errores gramaticales.	
Organización	<ul style="list-style-type: none">El documento está muy mal organizado.No se introduce bien el tema. El desarrollo no sigue una línea coherente. Se pasa de un aspecto a otro sin un orden. No hay un resumen y/o conclusiones.	<ul style="list-style-type: none">Hay algún aspecto claramente mejorable.La introducción no acaba de centrar bien el tema, o los diferentes apartados no acaban de estar bien ligados en una secuencia lógica, o falta un buen cierre con resumen y/o conclusiones.	<ul style="list-style-type: none">El documento está organizado de forma lógica.Las diferentes secciones y subsecciones están bien ligadas, y facilitan el seguimiento del contenido.El documento en una primera parte plantea el problema, en una segunda desarrolla y elabora la solución y en una tercera resume los resultados y plantea las conclusiones.	
Claridad	<ul style="list-style-type: none">El texto es muy difícil de entender; las frases son largas y confusas; constantemente se tienen que releer partes del texto para entenderlas y en varios casos finalmente no se entiende lo que se quiere decir.	<ul style="list-style-type: none">En alguna ocasión hay alguna frase larga y confusa que se tiene que releer varias veces para acabar de entender.	<ul style="list-style-type: none">Los contenidos son muy claros.Las frases son cortas y fáciles de entender a la primera.	
Uso de gráficos	<ul style="list-style-type: none">No se utilizan ni los gráficos, ni las figuras necesarias para facilitar la comprensión, o los que se utilizan no aclaran nada.	<ul style="list-style-type: none">En algún punto del documento se echa en falta algún gráfico o figura que ayude a aclarar los conceptos.Alguna de las figuras o gráficos parece poco clarificadora o innecesaria.	<ul style="list-style-type: none">Todos los gráficos y figuras utilizados tienen sentido y ayudan a entender la explicación.No hay ninguno que sobre, ni tampoco se echa ninguno en falta.	