

Diseño y Análisis de Algoritmos para el problema Gestión Agrícola

Amanda Cordero Lezcano
Christopher Guerra Herrero

Contents

1	Problema de Gestión Agrícola	2
1.1	Rotación de Cultivos en Ventanas Temporales	2
1.2	Asignación de Tareas	3
2	Problema Unbounded Knapsack	4
2.1	Demostración de que Unbounded Knapsack es NP-duro	4
2.2	Algoritmo exacto	6
2.2.1	Programación dinámica	6
2.2.2	Análisis del algoritmo	7
2.3	Algoritmo de aproximación	8
2.4	Experimentación	8
3	Problema Load Balancing	10
3.1	Demostración de que Load Balancing es NP duro	10
3.1.1	Demostración de que Partition es NP completo	11
3.2	Algoritmo Exacto	13
3.3	Algoritmos de aproximación	15
3.4	Experimentación	19

Chapter 1

Problema de Gestión Agrícola

La gestión agrícola moderna enfrenta desafíos críticos relacionados con la optimización de recursos, tanto en la planificación de cultivos como en la asignación eficiente de mano de obra. Estos problemas requieren enfoques algorítmicos robustos para maximizar beneficios económicos y minimizar tiempos de ejecución, asegurando un uso sostenible de los recursos disponibles. En este capítulo, se formalizan dos problemas centrales en este contexto y se establece su relación con problemas clásicos de optimización combinatoria.

1.1 Rotación de Cultivos en Ventanas Temporales

La mayoría de las fincas del país se especializan en un cultivo en particular. El tabaco, por ejemplo, ocupa el mayor por ciento de las tierras en Pinar del Río entre los meses de octubre a febrero (el período entre la siembra y la cosecha suele durar de 4 a 5 meses en dependencia de la variedad a sembrar y las condiciones climatológicas y del suelo).

El tabaco es una planta perteneciente a la familia de las solanáceas al igual que la papa, la berenjena y el tomate, estos son cultivos exigentes en nutrientes, y su crecimiento óptimo requiere un suelo fértil y bien equilibrado, rico en nitrógeno, fósforo y potasio principalmente.

Las solanáceas agotan rápidamente el suelo, por esta razón, es común realizar rotaciones de cultivos, ocupando el tiempo después de la cosecha de las mismas con otros cultivos que aporten nutrientes al suelo para disminuir la necesidad de fertilizantes de cara a la próxima siembra.

Formalización del problema

Se desean seleccionar cultivos para sembrar en un período determinado, considerando:

- **Aporte al suelo:** Cada cultivo aporta nutrientes al suelo en cierto grado.
- **Tiempo de ocupación:** Período durante el cual la tierra está ocupada (desde la siembra hasta la cosecha).
- **Ventana temporal fija:** Duración total disponible para la siembra.

Objetivo: Seleccionar un conjunto de cultivos (con posibles repeticiones) que maximice el beneficio total al suelo sin exceder la ventana temporal. Este problema es equivalente al **Problema de la Mochila Ilimitada (Unbounded Knapsack)**, donde:

- *Peso del elemento* \equiv Tiempo de ocupación del cultivo.
- *Valor del elemento* \equiv Beneficio que proporciona al suelo.
- *Capacidad de la mochila* \equiv Ventana temporal disponible.

1.2 Asignación de Tareas

Se debe distribuir un conjunto de tareas agrícolas (ej: podar, regar, cosechar) entre un grupo de obreros, considerando:

- **Tiempo de ejecución:** Cada tarea requiere un tiempo específico.
- **Recursos limitados:** Número fijo de obreros disponibles.

Objetivo: Asignar tareas a obreros minimizando el tiempo total de ejecución (*makespan*). Este problema se modela como **Balanceo de Carga (Load Balancing)** en sistemas distribuidos.

Chapter 2

Problema Unbounded Knapsack

El problema **Unbounded Knapsack (UK)** es una extensión del clásico problema de la mochila (**Knapsack (K)**), donde los elementos pueden ser seleccionados múltiples veces. Este problema tiene aplicaciones en diversas áreas, como la logística, la planificación de recursos y la optimización de inventarios. En este texto, exploraremos la relación entre **UK** y **K**, y demostraremos que **UK** es un problema NP-duro mediante una reducción formal desde **UK** a **K**.

2.1 Demostración de que Unbounded Knapsack es NP-duro

Para demostrar que el problema **Unbounded Knapsack (UK)** es NP-duro, realizaremos una reducción desde **UK** a **Knapsack (K)**, que es un problema NP-completo. La idea central es transformar una instancia de **UK** en una instancia equivalente de **K**, resolviendo subproblemas de manera iterativa hasta obtener una solución completa.

Definición del problema UK

Dada una instancia de **UK**, se tiene:

- Un conjunto de n elementos, donde cada elemento i tiene un peso w_i y un valor v_i .
- Una capacidad máxima de la mochila W .

El objetivo es maximizar el valor total de los elementos seleccionados sin exceder la capacidad W , permitiendo que cada elemento pueda ser seleccionado múltiples veces.

Reducción de UK a K

La reducción se realiza mediante el algoritmo siguiente:

Algorithm 1 Reducción de UK a K

1. Inicialización:

$S \leftarrow \emptyset$ {Conjunto de elementos seleccionados}

$P \leftarrow 0$ {Peso total acumulado}

$R \leftarrow W$ {Capacidad residual inicial}

2. Iteración:

while $R > 0$ **do**

a. Resolver **UK** sobre el conjunto de elementos E con capacidad R .

b. Sea S_{temp} el conjunto de elementos seleccionados (sin repeticiones) y P_{temp} el peso total de S_{temp} .

if $S_{\text{temp}} = \emptyset$ **then**

Break {No se pueden agregar más elementos}

end if

c. Actualizar:

$S \leftarrow S \cup S_{\text{temp}}$

$P \leftarrow P + P_{\text{temp}}$

$R \leftarrow R - P_{\text{temp}}$

d. Eliminar los elementos de S_{temp} de E : $E \leftarrow E \setminus S_{\text{temp}}$

end while

3. Terminación:

Devolver S {Conjunto de elementos seleccionados}

Demostración de correctitud

Para demostrar que la reducción de **UK** a **K** es válida, debemos probar que la solución obtenida mediante el algoritmo descrito es equivalente a la solución óptima de **K**. A continuación, se detallan los aspectos clave de la demostración:

- **No repetición de elementos:** En cada iteración, el algoritmo selecciona un subconjunto de elementos **sin repeticiones**, es decir, cada elemento se incluye en la mochila a lo sumo una vez. Esto es consistente con la definición de **K**, donde los elementos no pueden repetirse.

- **Optimalidad:** Los elementos sin repetición de cada iteración de **UK** son parte de la solución óptima de **K** ya que si alguno no lo fuera es porque este puede ser remplazado por otro u otros que en total maximizen la ganancia, pero en tal caso **UK** los hubiera seleccionados.
- **Conjunto finito de elementos:** Dado que el conjunto de elementos es finito, el algoritmo termina en un número finito de pasos. En cada iteración, se elimina al menos un elemento del conjunto de candidatos, lo que garantiza que el proceso no continúe indefinidamente.
- **Complejidad de la transformación:** La reducción realiza una transformación en tiempo polinomial. En el peor caso, el algoritmo ejecuta **UK** una vez por cada elemento distinto, lo que resulta en una complejidad de $O(n \cdot T_{UK})$, donde T_{UK} es la complejidad de resolver **UK**. Si **UK** fuera resoluble en tiempo polinomial, entonces **K** también lo sería, ya que la transformación preserva la clase de complejidad.
- **Uso completo de la capacidad:** El algoritmo termina cuando no se pueden agregar más elementos sin exceder la capacidad de la mochila. Esto asegura que la capacidad W se utiliza de manera óptima, y la solución acumulada es equivalente a la solución óptima de **K**.

Mediante esta reducción, hemos demostrado que cualquier instancia de **K** puede ser transformada en una secuencia de tamaño polinomial de instancias de **UK**. Dado que **K** es un problema NP-completo, concluimos que **UK** es NP-duro.

2.2 Algoritmo exacto

Ataquemos el problema con Programación dinámica. La idea de solución es similar a la vista en [1] para el problema 0/1 Knapsack. Utilizamos un arreglo DP donde $DP[i]$ almacena el valor máximo alcanzable con una mochila de capacidad i . Comenzamos con capacidad 0 y progresivamente resolvemos capacidades mayores usando soluciones ya calculadas.

Subestructura óptima

La propiedad clave es que la solución óptima para capacidad i puede construirse a partir de soluciones óptimas para capacidades menores. Si un elemento de peso w_j se incluye en la solución óptima para capacidad i , entonces el valor óptimo será $v_j + DP[i - w_j]$.

2.2.1 Programación dinámica

Estados y transiciones

- **Estado:** $DP[i]$ representa el valor máximo para capacidad i .
- **Inicialización:** $DP[0] = 0$ (ningún elemento cabe en capacidad 0).

- **Transición:** Para cada capacidad i y elemento j :

$$DP[i] = \max(DP[i], v_j + DP[i - w_j]) \quad \text{si } w_j \leq i$$

Algorithm 2 Pseudocódigo Unbounded Knapsack

$W, \text{weights}, \text{values}$

Inicializar arreglo $DP[0 \dots W]$ con 0's

for $i = 1$ to W **do**

for $j = 0$ to $n - 1$ **do**

if $\text{weights}[j] \leq i$ **then**

$DP[i] \leftarrow \max(DP[i], DP[i - \text{weights}[j]] + \text{values}[j])$

end if

end for

end for

return $DP[W]$

Complejidad

- **Temporal:** $O(n \cdot W)$ (dos bucles anidados: W iteraciones externas, n internas).
- **Espacial:** $O(W)$ (solo se almacena un arreglo de tamaño $W + 1$).

2.2.2 Análisis del algoritmo

Correctitud

La demostración se basa en inducción matemática:

- **Caso base:** $DP[0] = 0$ es trivialmente correcto.
- **Hipótesis inductiva:** Supongamos que $DP[k]$ es óptimo para todo $k < i$.
- **Paso inductivo:** Para capacidad i , consideramos todos los elementos j con $w_j \leq i$. Si el elemento j es parte de la solución óptima, entonces:

$$DP[i] = v_j + DP[i - w_j]$$

Por hipótesis inductiva, $DP[i - w_j]$ ya es óptimo. Al maximizar sobre todos los j , garantizamos que $DP[i]$ también es óptimo.

El procesamiento en orden ascendente de capacidades asegura que al calcular $DP[i]$, todos los valores $DP[k]$ para $k < i$ ya han sido computados correctamente, permitiendo reutilizar soluciones previas.

2.3 Algoritmo de aproximación

Con información de [2]

Comencemos por describir un algoritmo que garantiza una solución dentro de un factor de 2 del valor óptimo. Supongamos que el artículo j tiene la mayor densidad de beneficio, definida como p_i/w_i , donde p_i es el beneficio del artículo i y w_i es su peso. Denotamos z_L como el valor objetivo obtenido al seleccionar tantas unidades del artículo j como sea posible sin exceder la capacidad de la mochila C . Formalmente:

$$z_L = p_j \lfloor C/w_j \rfloor,$$

donde $\lfloor C/w_j \rfloor$ representa el número máximo de unidades enteras del artículo j que caben en la mochila.

Es importante notar que si se relaja la restricción de integridad (es decir, si se permite fraccionar los artículos), el valor objetivo óptimo sería:

$$z = p_j(C/w_j),$$

el cual constituye una cota superior para el valor óptimo z^* bajo las restricciones originales. Por lo tanto, se cumple que:

$$z_L \leq z^* \leq z.$$

Además, dado que $z = z_L + p_j\epsilon$, donde $0 \leq \epsilon < 1$, se deduce que:

$$z \leq 2z_L.$$

En consecuencia, tenemos que:

$$z^* \leq 2z_L,$$

lo cual implica que el algoritmo que selecciona únicamente el artículo con mayor densidad de beneficio y llena la mochila con este tipo de elementos constituye una 2-aproximación al valor óptimo del problema.

2.4 Experimentación

La Figura 2.1 compara el valor máximo obtenido por el algoritmo exacto (programación dinámica) y el algoritmo de 2-aproximación para distintas instancias del problema Unbounded Knapsack.

El algoritmo de 2-aproximación, aunque subóptimo, nunca excede el doble del valor óptimo, confirmando su garantía teórica

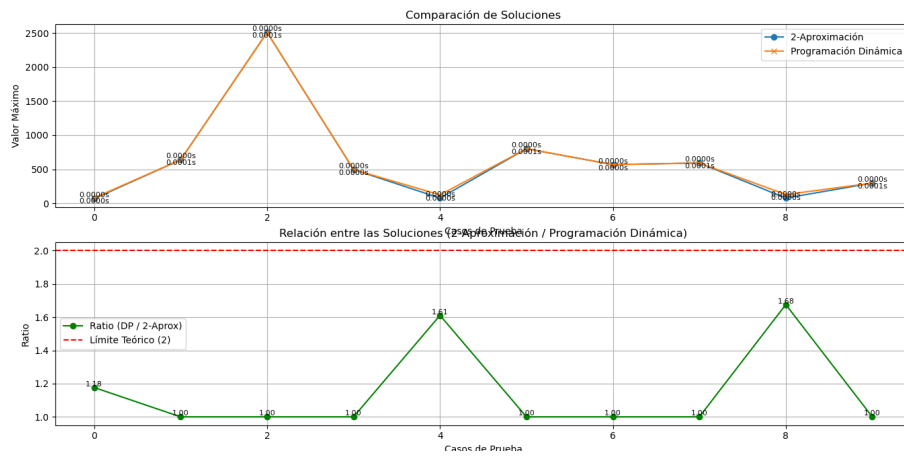


Figure 2.1: Comparación gráfica entre los algoritmos

Chapter 3

Problema Load Balancing

Con información de [3] y [4]

Formulamos el **Problema de Load Balancing** de la siguiente manera. Se nos da un conjunto de m máquinas M_1, \dots, M_m y un conjunto de n trabajos; cada trabajo j tiene un tiempo de procesamiento t_j . Buscamos asignar cada trabajo a una de las máquinas de manera que las cargas colocadas en todas las máquinas estén lo más *balanceadas* posible.

Más concretamente, en cualquier asignación de trabajos a máquinas, podemos denotar por $A(i)$ el conjunto de trabajos asignados a la máquina M_i . Bajo esta asignación, la máquina M_i trabajará un tiempo total de

$$T_i = \sum_{j \in A(i)} t_j,$$

y declaramos que esta es la carga en la máquina M_i . Buscamos minimizar una cantidad conocida como el *makespan*, que es simplemente la carga máxima en cualquier máquina:

$$T = \max_i T_i.$$

Nuestro problema de optimización tiene una versión de decisión: Dado un valor M , ¿existe una asignación de tareas tal que el makespan sea $\leq M$? A continuación nos apoyaremos del problema de decisión y del conocido problema Partition para demostrar que la versión de optimización de Load Balancing es NP duro.

3.1 Demostración de que Load Balancing es NP duro

Para demostrar que el problema de decisión es NP-completo, realizamos una reducción desde el problema **Partition**, primero demostremos que este es NP-

completo reduciéndolo a Subset-Sum.

3.1.1 Demostración de que Partition es NP completo

Problema Partition: Dado un conjunto de números $\{p_1, p_2, \dots, p_n\}$, ¿es posible dividirlo en dos subconjuntos disjuntos A y B tales que:

$$\sum_{p_i \in A} p_i = \sum_{p_i \in B} p_i?$$

Sea $P = \sum_{i=1}^n p_i$. El problema puede reformularse como: ¿existe una partición tal que:

$$\sum_{p_i \in A} p_i = \frac{P}{2}?$$

Para demostrar que **Partition** es NP-completo, seguimos los siguientes pasos:

Paso 1: Partition está en NP

Un problema está en NP si, dada una solución candidata, podemos verificar su validez en tiempo polinomial.

Una solución candidata para **Partition** consiste en un subconjunto $A \subseteq S$. Para verificar si A es una solución válida:

- Calculamos $\text{suma}(A) = \sum_{x \in A} x$.
- Verificamos si $\text{suma}(A) = \frac{\sum_{i=1}^n a_i}{2}$.
- Esto se puede hacer en tiempo polinomial con respecto al tamaño de S .

Por lo tanto, **Partition** está en NP.

Paso 2: Partition es NP-duro

Para demostrar que **Partition** es NP-duro, realizamos una reducción polinomial desde el problema **Subset-Sum**, que es conocido como NP-completo.

Problema Subset-Sum

El problema **Subset-Sum** se define como sigue:

Entrada: Un conjunto finito $S = \{a_1, a_2, \dots, a_n\}$ de números enteros positivos y un objetivo T .

Pregunta: ¿Existe un subconjunto $A \subseteq S$ tal que:

$$\sum_{x \in A} x = T?$$

Reducción desde Subset-Sum a Partition

Dada una instancia del problema **Subset-Sum** con conjunto $S = \{a_1, a_2, \dots, a_n\}$ y objetivo T , construimos una instancia del problema **Partition** como sigue:

1. Calcula la suma total de los elementos de S :

$$P = \sum_{i=1}^n a_i.$$

2. Agrega un nuevo número b al conjunto S , donde:

$$b = |P - 2T|.$$

3. Define el nuevo conjunto $S' = S \cup \{b\}$.

Ahora preguntamos: ¿Es posible particionar S' en dos subconjuntos A y B tales que:

$$\sum_{x \in A} x = \sum_{y \in B} y?$$

En tal caso la cardinalidad de los subconjuntos será

$$\frac{P + |P - 2T|}{2}$$

Por lo que uno de los dos tendrá T de cardinalidad al excluir a b

Reducción desde Partition a Load Balancing

Dada una instancia del problema Load Balancing de decisión, construimos el problema Partition como sigue:

- Sea $m = 2$ (dos máquinas).
- Asigna los tiempos de procesamiento p_1, p_2, \dots, p_n a las tareas.
- Define $T = \frac{P}{2}$.

Ahora preguntamos: ¿Existe una asignación de tareas a las dos máquinas tal que el makespan sea $\leq T$?

Equivalencia

Si existe una asignación de tareas entonces habrá una partición. Si no existe, no existirá la partición. Por lo tanto, resolver el problema de decisión de Load Balancing permite resolver el problema Partition. Como el problema Partition es NP-completo y hemos mostrado que cualquier instancia de Partition puede reducirse en tiempo polinomial a una instancia del problema de decisión de Load Balancing, y la solución de este es verificable en tiempo polinomial de forma trivial (pertenece a NP), concluimos que el problema de decisión de Load Balancing es NP-completo.

Finalmente, demostramos que el problema de optimización de Load Balancing es NP-duro. Para ello, notamos que:

1. Si podemos resolver el problema de optimización (encontrar el makespan mínimo T^*), entonces podemos resolver el problema de decisión para cualquier valor T comparando T con T^* :

$$\text{Respuesta al problema de decisión} = \begin{cases} \text{sí,} & \text{si } T \geq T^*, \\ \text{no,} & \text{si } T < T^*. \end{cases}$$

2. Como el problema de decisión es NP-completo, resolver el problema de optimización implica resolver un problema NP-completo.

Por lo tanto, el problema de optimización de Load Balancing es NP-duro.

3.2 Algoritmo Exacto

Búsqueda con Backtracking y Poda es una alternativa para resolver el problema de Load Balancing, aunque su escalabilidad está limitada por la complejidad computacional del problema.

Esta técnica explora sistemáticamente todas las posibles asignaciones de trabajos a máquinas, pero utiliza estrategias de poda para eliminar ramas del árbol de búsqueda que no pueden conducir a una solución mejor que la ya encontrada.

Pasos Clave

1. Ordenar Trabajos

- Ordenar los trabajos en orden descendente de t_j .
- **Propósito:** Asignar primero los trabajos más grandes para revelar rápidamente soluciones subóptimas y permitir una poda temprana.

2. Inicializar Variables

- Mantener dos registros fundamentales:
 1. T_{best} : El mejor makespan encontrado hasta el momento (inicializado con una cota superior, como la suma total de t_j).
 2. $\text{cargas} = [c_1, c_2, \dots, c_m]$: Arreglo que almacena la carga parcial de cada máquina M_i (inicializado en 0).

3. Exploración Recursiva

- Para cada trabajo j (en orden descendente):
 1. Asignar j a cada máquina M_i .
 2. Actualizar la carga parcial de M_i : $c_i \leftarrow c_i + t_j$.

3. Si en cualquier momento $c_i \geq T_{\text{best}}$, detener la exploración de esa rama (**poda por optimalidad**).
4. Llamar recursivamente al algoritmo para el siguiente trabajo $j + 1$.
5. Deshacer la asignación ($c_i \leftarrow c_i - t_j$) para probar otras máquinas (**backtracking**).

4. Estrategias de Poda

- **Poda por Carga Parcial:** Si $c_i > T_{\text{best}}$, ignorar asignaciones futuras a M_i .
- **Cota Inferior:** Calcular una cota inferior para el makespan parcial usando:

$$T_{\min} = \max \left(\max_{1 \leq i \leq m} c_i, \frac{1}{m} \left(\sum_{i=1}^m c_i + \sum_{j=k}^n t_j \right) \right)$$

donde k es el índice del trabajo actual. Si $T_{\min} \geq T_{\text{best}}$, podar la rama.

Complejidad Temporal

$O(m^n)$ en el peor caso

Algorithm 3 Backtracking con Poda para Load Balancing

Lista de trabajos $jobs$, número de máquinas m

Makespan mínimo T_{best}

Ordenar $jobs$ en orden descendente de t_j $\{t_1 \geq t_2 \geq \dots \geq t_n\}$

Inicializar $T_{\text{best}} \leftarrow \sum_{j=1}^n t_j$ {Cota superior inicial}

Inicializar cargas $loads \leftarrow [0, 0, \dots, 0]$ {Vector de tamaño m }

function BACKTRACK($current_job$, $loads$, T_{best})

if $current_job = n$ **then**

return $\max(load)$ {Base case: todos los trabajos asignados}

end if

$job \leftarrow jobs[current_job]$

for $i \leftarrow 1$ **to** m **do**

if $loads[i] + job \geq T_{\text{best}}$ **then**

 {Poda: rama no prometedora}

end if

$loads[i] \leftarrow loads[i] + job$ {Asignar trabajo a máquina i }

$new_T \leftarrow \text{BACKTRACK}(current_job + 1, loads, T_{\text{best}})$

if $new_T < T_{\text{best}}$ **then**

$T_{\text{best}} \leftarrow new_T$ {Actualizar mejor solución}

end if

$loads[i] \leftarrow loads[i] - job$ {Backtracking: deshacer asignación}

end for

return T_{best}

end function

return BACKTRACK(0, $loads$, T_{best}) {Llamada inicial}

3.3 Algoritmos de aproximación

Hemos demostrado que el problema de Load Balancing (en su versión de optimización) es NP-duro mediante una reducción desde el problema Partition al problema de decisión, y luego conectando el problema de decisión con el problema de optimización. Ahora vamos a ver algunos algoritmos para aproximar la solución del mismo.

Greedy-Balance

Primero consideramos un algoritmo greedy muy simple para el problema. El algoritmo hace un recorrido por los trabajos en cualquier orden; cuando llega al trabajo j , lo asigna a la máquina cuya carga sea la más pequeña hasta ese momento.

Algorithm 4 Greedy-Balance

Empezar sin trabajos asignados

for cada máquina M_i **do**

$T_i \leftarrow 0$ {Inicializar la carga de la máquina}

$A(i) \leftarrow \emptyset$ {Inicializar el conjunto de trabajos asignados a M_i }

end for

for cada trabajo $j = 1, \dots, n$ **do**

 Sea M_i la máquina con la menor carga

 Asignar el trabajo j a la máquina M_i

$A(i) \leftarrow A(i) \cup \{j\}$

$T_i \leftarrow T_i + t_j$

end for

Análisis del Algoritmo

Sea T el makespan de la asignación resultante; queremos mostrar que T no es mucho mayor que el makespan mínimo posible T^* . Por supuesto, al intentar hacer esto, nos encontramos inmediatamente con el problema básico mencionado anteriormente: necesitamos comparar nuestra solución con el valor óptimo T^* , aunque no sabemos cuál es este valor y no tenemos forma de calcularlo. Para el análisis, por lo tanto, necesitaremos una cota inferior para el óptimo, una cantidad con la garantía de que, no importa cuán bueno sea el óptimo, no puede ser menor que esta cota.

Hay muchas posibles cotas inferiores para el óptimo. Una idea para una cota inferior se basa en considerar el tiempo total de procesamiento $\sum_j t_j$. Una de las m máquinas debe hacer al menos una fracción $1/m$ del trabajo total, y por lo tanto tenemos lo siguiente:

$$T^* \geq \frac{1}{m} \sum_j t_j. \quad (3.1)$$

Hay un tipo particular de caso en el que esta cota inferior es demasiado débil para ser útil. Supongamos que tenemos un trabajo que es extremadamente largo en relación con la suma de todos los tiempos de procesamiento. En una versión suficientemente extrema de esto, la solución óptima colocará este trabajo en una máquina por sí solo, y será el último en terminar. En tal caso, nuestro algoritmo

greedy en realidad produciría la solución óptima, pero la cota inferior en (3.1) no es lo suficientemente fuerte para establecer esto.

Esto sugiere la siguiente cota inferior adicional para T^* :

$$T^* \geq \max_j t_j. \quad (3.2)$$

Ahora estamos listos para evaluar la asignación obtenida por nuestro algoritmo greedy.

El algoritmo **Greedy-Balance** produce una asignación de trabajos a máquinas con un makespan $T \leq 2T^*$.

Aquí está el plan general para la demostración. Al analizar un algoritmo de aproximación, se compara la solución obtenida con lo que se sabe sobre el óptimo; en este caso, nuestras cotas inferiores (3.1) y (3.2). Consideramos una máquina M_i que alcanza la carga máxima T en nuestra asignación, y nos preguntamos: ¿Cuál fue el último trabajo j asignado a M_i ? Si t_j no es demasiado grande en relación con la mayoría de los otros trabajos, entonces no estamos muy por encima de la cota inferior (3.1). Y, si t_j es un trabajo muy grande, entonces podemos usar (3.2).

Cuando asignamos el trabajo j a M_i , la máquina M_i tenía la carga más pequeña de cualquier máquina; esta es la propiedad clave de nuestro algoritmo greedy. Su carga justo antes de esta asignación era $T_i - t_j$, y como esta era la carga más pequeña en ese momento, se sigue que todas las máquinas tenían una carga de al menos $T_i - t_j$. Así, sumando las cargas de todas las máquinas, tenemos:

$$\sum_k T_k \geq m(T_i - t_j),$$

o equivalentemente,

$$T_i - t_j \leq \frac{1}{m} \sum_k T_k.$$

Pero el valor $\sum_k T_k$ es simplemente la carga total de todos los trabajos $\sum_j t_j$ (ya que cada trabajo se asigna exactamente a una máquina), y por lo tanto la cantidad en el lado derecho de esta desigualdad es exactamente nuestra cota inferior en el valor óptimo, de (3.1). Así,

$$T_i - t_j \leq T^*.$$

Ahora consideramos la parte restante de la carga en M_i , que es solo el trabajo final j . Aquí simplemente usamos la otra cota inferior que tenemos, (3.2), que dice que $t_j \leq T^*$. Sumando estas dos desigualdades, vemos que:

$$T_i = (T_i - t_j) + t_j \leq 2T^*.$$

Como nuestro makespan T es igual a T_i , este es el resultado que queremos.

Sorted-Balance

Ahora pensemos en cómo podríamos desarrollar un algoritmo de aproximación mejor, es decir, uno para el cual siempre estemos garantizados de estar dentro de un factor estrictamente menor que 2 del óptimo. Para hacer esto, es útil pensar en los peores casos para nuestro algoritmo de aproximación actual.

Un ejemplo malo para nuestro algoritmo greedy pudiera tener la siguiente característica: distribuimos todo de manera muy uniforme entre las máquinas, y luego llega un último trabajo gigante. Intuitivamente, parece que ayudaría organizar primero los trabajos más grandes de manera adecuada, con la idea de que, más tarde, los trabajos pequeños solo pueden causar un daño limitado. Y, de hecho, esta idea conduce a una mejora cuantificable.

Así, ahora analizamos la variante del algoritmo greedy que primero ordena los trabajos en orden decreciente de tiempo de procesamiento y luego procede como antes. Demostraremos que la asignación resultante tiene un makespan que es, como máximo, 1.5 veces el óptimo.

Algorithm 5 Sorted-Balance

Iniciar sin trabajos asignados.

Asignar $T_i = 0$ y $A(i) = \emptyset$ para todas las máquinas M_i .

Ordenar los trabajos en orden decreciente de tiempos de procesamiento t_j .

Suponer que $t_1 \geq t_2 \geq \dots \geq t_n$.

for cada trabajo $j = 1, \dots, n$ **do**

 Sea M_i la máquina que alcanza el mínimo $\min_k T_k$.

 Asignar el trabajo j a la máquina M_i .

 Actualizar $A(i) \leftarrow A(i) \cup \{j\}$.

 Actualizar $T_i \leftarrow T_i + t_j$.

end for

La mejora proviene de la siguiente observación. Si tenemos menos de m trabajos, entonces la solución greedy claramente será óptima, ya que coloca cada trabajo en su propia máquina. Y si tenemos más de m trabajos, entonces podemos usar la siguiente cota inferior adicional para el óptimo.

$$T^* \geq 2t_{m+1}. \quad (3.3)$$

Considera solo los primeros $m+1$ trabajos en el orden clasificado. Cada uno de ellos toma al menos un tiempo t_{m+1} . Hay $m+1$ trabajos y solo m máquinas, por lo que debe haber una máquina a la que se le asignen dos de estos trabajos. Esta máquina tendrá un tiempo de procesamiento de al menos $2t_{m+1}$.

El algoritmo **Sorted-Balance** produce una asignación de trabajos a máquinas con un makespan $T \leq \frac{3}{2}T^*$.

La demostración será muy similar al análisis del algoritmo anterior. Como antes, consideraremos una máquina M_i que tiene la carga máxima. Si M_i solo contiene un trabajo, entonces el programa es óptimo.

Supongamos que la máquina M_i tiene al menos dos trabajos, y sea t_j el último trabajo asignado a la máquina. Nótese que $j \geq m+1$, ya que el algoritmo asignará los primeros m trabajos a m máquinas distintas. Por lo tanto, $t_j \leq t_{m+1} \leq \frac{1}{2}T^*$, donde la segunda desigualdad es (3.3).

Ahora procedemos como anteriormente, con el siguiente cambio. Al final de la demostración, teníamos las desigualdades $T_i - t_j \leq T^*$ y $t_j \leq T^*$, y las sumamos para obtener el factor de 2. Pero en nuestro caso aquí, la segunda de estas desigualdades es, de hecho, $t_j \leq \frac{1}{2}T^*$; por lo tanto, sumar las dos desigualdades nos da la cota:

$$T_i \leq \frac{3}{2}T^*.$$

3.4 Experimentación

La Figura 3.1 contrasta el makespan del algoritmo exacto (backtracking con poda) y el algoritmo Greedy-Balance (2-aproximación). En todos los casos, el makespan de Greedy-Balance se mantiene por debajo de $2T^*$, incluso en distribuciones heterogéneas donde un trabajo grande domina el makespan.

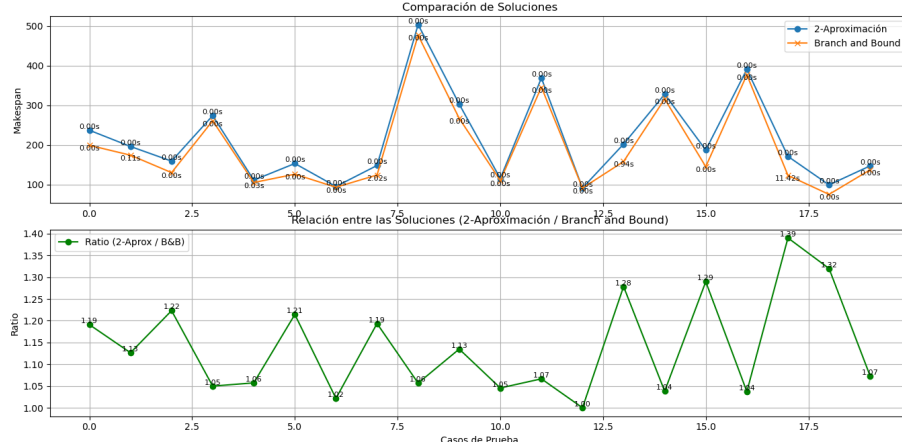


Figure 3.1: Comparación gráfica entre los algoritmos

Bibliography

- [1] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. 3rd ed. MIT Press, 2009.
- [2] Donguk Rhee. *Esquemas de Aproximación Polinomial Más Rápidos para Problemas de la Mochila*. Tesis de Maestría en Investigación de Operaciones, Instituto Tecnológico de Massachusetts, 2016.
- [3] Jon Kleinberg and Éva Tardos. *Algorithm Design*. 1st ed. Pearson, 2005. Extraído de página 600.
- [4] Michael R. Garey and David S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness*. 1st ed. W. H. Freeman and Company, 1979. Extraído de página 238.