Analyzing Recursive Algorithms (Ch. 4)

A recursive algorithm can often be described by a recurrence equation that describes the overall runtime on a problem of size \( n \) in terms of the runtime on smaller inputs.

For divide-and-conquer algorithms, we get recurrences like:

\[
T(n) = \begin{cases} 
1 & \text{if } n = c \\
\sum_{i=0}^{\log n} aT(n/b^i) + D(n) + C(n) & \text{otherwise}
\end{cases}
\]

where

- \( a \) = number of subproblems we divide the problem into
- \( n/b \) = size of the subproblems (in terms of \( n \))
- \( D(n) \) = time to divide the size \( n \) problem into subproblems
- \( C(n) \) = time to combine the subproblem solutions to get the answer for the problem of size \( n \)

Solving Recurrences

We will use the following methods to solve recurrences

1. Backward Substitution.
2. Apply the “Master Theorem”: If the recurrence has the form

\[
T(n) = aT(n/b) + f(n)
\]

then there are 2 formulae that can (often) be applied; one of these is given in § 4-3.

Recurrence trees can be used along with backward substitution to guess the running time of a recurrence relation. Most recurrences of the form

\[
T(n) = aT(n/b) + f(n)
\]

will be solved using the Master Theorem.

To make the solutions simpler, we will

- assume base cases are constant, i.e., \( T(n) = \Theta(1) \) for \( n \) small enough.

Solving recurrence with backward substitution

Algorithm \( F(n) \)

Input: a positive integer \( n \)
Output: \( n! \)
1. if \( n = 0 \) return 1
2. else return \( F(n-1) \times n \)

\[ T(n) = \begin{cases} 
T(n-1) + 1 & \text{if } n \leq 1 \\
1T(n-1) + D(n) + C(n) & \text{otherwise}
\end{cases} \]

We solved this recurrence (i.e., found an expression of the running time \( T(n) \) that is not given in terms of itself) using a method known as backward substitution.

Analyzing Recursive Algorithms

For recursive algorithms such as computing the factorial of \( n \), we get an expression like the following:

\[
T(n) = \begin{cases} 
1 & \text{if } n = 0 \\
1T(n-1) + D(n) + C(n) & \text{otherwise}
\end{cases}
\]

where

- \( a \) = number of subproblems (1)
- \( n-1 \) = size of the subproblems (in terms of \( n \))
- \( D(n) \) = time to divide the size \( n \) problem into subproblems = 1
- \( C(n) \) = time to combine the subproblem solutions to get the answer for the problem of size \( n = 1 \)

Solving Recurrences: Backward Substitution

Example: \( T(n) = 2T(n/2) + n \)

\[
T(n) = 2T(n/2) + n
= 2[2T(n/4) + n/2] + n
= 4[2T(n/4) + n/4] + n/2 + n
= 8[2T(n/4) + n/8] + n/4 + n/2 + n
= 16[2T(n/4) + n/16] + n/8 + n/4 + n/2 + n
= \ldots
\]

\[
\ldots \text{continue until } T(n/2) \text{ is reached}
\]

\[
= 2^{\log_2 n}[2(n/2^n) + \ldots + n/2^n + n/2^{n-1} + n/2^{n-2} + \ldots + n/2^1 + n/2^0 ]
= 2n + n \log n
= O(n \log n)
\]

Recursion Tree for \( T(n) = 2T(n/2) + n \)

\[
T(n) = \begin{cases} 
c & \text{if } n = 1 \\
2T(n/2) + n & \text{otherwise}
\end{cases}
\]
Solving Recurrences: Backward Substitution

Example:

\[ T(n) = 2T(n/2) + 4n \]

- \[ T(n) = 2T(n/2) + 4n \]
- \[ T(n) = 2T(n/4) + 4n/2 + 4n \]
- \[ T(n) = 4T(n/4) + 4n + 4n \]
- \[ T(n) = 8T(n/8) + 16n/4 + 4n + 4n \]
- \[ \ldots \]
- \[ T(n) = 2^{\lg n} T(1) + \ldots + 4n + 4n + 4n \]
- \[ T(n) = c \cdot 2^{\lg n} + n (\text{after } \lg n \text{ iterations}) \]
- \[ T(n) = O(n^{\lg 2}) \]

Solving Recurrences: Backward Substitution

Example:

\[ T(n) = 4T(n/2) + n \]

- \[ T(n) = 4T(n/2) + n \]
- \[ T(n) = 16T(n/4) + 4n/2 + n \]
- \[ T(n) = 64T(n/8) + 16n/4 + 2n + n \]
- \[ \ldots \]
- \[ T(n) = c \cdot 2^{\lg n} + n (\text{after } \lg n \text{ iterations}) \]
- \[ T(n) = O(n^{\lg 2}) \]

Binary Search (recursive version)

**Algorithm** Binary-Search-Rec(A[1…n], k, l, r)

Input: a sorted array A of n comparable items, search key k, leftmost and rightmost index positions in A

Output: Index of array's element that is equal to k or -1 if k not found

1. if (l > r) return -1; k not found
2. else
3. \[ m = \lfloor (l + r)/2 \rfloor \]; m is midpoint
4. if \( k = A[m] \) return m; found index of k at A(n)
5. else if \( k < A[m] \) return Binary-Search-Rec(A[k, m-1], l); look in lower half
6. else return Binary-Search-Rec(A[k, m+1, r]); look in upper half

What is the running time of this algorithm for an input of size n? We need to figure out what \( T(n) \) is for this algorithm.

Analysis of Divide-and-Conquer Algorithms

The divide-and-conquer paradigm (Ch.3)

- `divide` the problem into a number of subproblems
- `conquer` the subproblems (solve them)
- `combine` the subproblem solutions to get the solution to the original problem

**Example: Merge Sort**

- `divide` the n-element sequence to be sorted into two n/2-element sequences.
- `conquer` the subproblems recursively using merge sort.
- `combine` the resulting two sorted n/2-element sequences by merging.

Analyzing Merge-Sort

**Divide** (lgn levels)

<table>
<thead>
<tr>
<th>2</th>
<th>8</th>
<th>1</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>7</th>
<th>6</th>
</tr>
</thead>
</table>

**Merge**

<table>
<thead>
<tr>
<th>2</th>
<th>8</th>
<th>1</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>7</th>
<th>6</th>
</tr>
</thead>
</table>
Analyzing Merge-Sort

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>C(n)</td>
<td>Θ(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Recurrence for worst-case running time for Merge-Sort

\[ T(n) = \begin{cases} 
\Theta(1) & \text{if } n = 1 \\
2T(n/2) + \Theta(n) & \text{otherwise} 
\end{cases} \]

The Merge subroutine takes \( \Theta(n) \) time to merge \( n \) elements that are divided into two sorted arrays of \( n/2 \) elements each.

Solving Recurrences: Master Method (§4.3)

The master method provides a 'cookbook' method for solving recurrences of a certain form.

**Master Theorem:** Let \( a \geq 1 \) and \( b > 1 \) be constants, let \( f(n) \) be a function, and let \( T(n) \) be defined on nonnegative integers as:

\[ T(n) = aT(n/b) + f(n) \]

Where \( a \) is the number of subproblems, \( n/b \) is the size of each subproblem, and \( f(n) \) is the time to divide or combine data.

Then, \( T(n) \) can be bounded asymptotically as follows:

1. \( T(n) = \Theta(n^{\log_b a}) \) if \( f(n) = O(n^{\log_b a-\epsilon}) \) for some constant \( \epsilon > 0 \)
2. \( T(n) = \Theta(n^{\log_b a} \log n) \) if \( f(n) = \Theta(n^{\log_b a}) \)
3. \( T(n) = \Theta(f(n)) \) if \( f(n) = \omega(n^{\log_b a}) \) for some constant \( \epsilon > 0 \)

What is \( \log_b a? \) The number of times we divide a by \( b \) to reach \( O(1) \).
Solving Recurrences: Master Method (§4.3)

**Master Theorem**: Let \( a \geq 1 \) and \( b > 1 \) be constants, let \( f(n) \) be a function, and let \( T(n) \) be defined on nonnegative integers as:

\[
T(n) = aT(n/b) + f(n)
\]

Then, \( T(n) \) can be bounded asymptotically as follows:

1. \( T(n) = \Theta(n^\log_b a) \) if \( f(n) = O(n^{\log_b a - \epsilon}) \) for some constant \( \epsilon > 0 \)
2. \( T(n) = \Theta(n^\log_b a \log^{k+1} n) \) if \( f(n) = \Omega(n^{\log_b a - \epsilon}) \) for some constant \( \epsilon > 0 \)
3. \( T(n) = O(n^\log_b a) \) if \( f(n) = \omega(n^{\log_b a}) \)

Why? The proof uses a recursion tree argument (given in our textbook).

Solving Recurrences: Master Method (§4.3)

A more general version of Case 2 follows:

\[
T(n) = \Theta(n^\log_b a \log^{k+1} n) \quad \text{if} \quad f(n) = O(n^{\log_b a \log^{k+1} n}) \quad \text{for} \quad k \geq 0
\]

This case covers the gap between cases 2 and 3 in which \( f(n) \) is larger than \( n^{\log_b a} \) by only a polylog factor. We'll see an example of this type of recurrence in class.

Alternate Version of Master Method

**Master Theorem**: Let \( a \geq 1 \), \( b > 1 \), \( k \geq 0 \) be constants, let \( f(n) \) be a function, and let \( T(n) \) be defined on nonnegative integers as:

\[
T(n) = aT(n/b) + \Theta(n^k \log^m n)
\]

Then, \( T(n) \) can be bounded asymptotically as follows:

1. \( a > b^k \), then \( T(n) = \Theta(n^k \log^{m+1} n) \)
2. \( a = b^k \), then \( T(n) = \Theta(n^k \log^{m+1} n) \)
3. \( a < b^k \), then \( T(n) = \Theta(n^k \log^m n) \)

Solving Recurrences: Master Method

Example: \( T(n) = 9T(n/3) + n \)

- \( a = 9, b = 3, f(n) = n \), \( n^{\log_b a} = n^{\log_3 9} = n^2 \)
- compare \( f(n) = n \) with \( n^2 \)
- \( n = O(n^{k+1}) \) (so \( f(n) \) is polynomially smaller than \( n^{\log_b a} \))
- case 1 applies: \( T(n) \in \Theta(n^2) \)

Example: \( T(n) = T(n/2) + 1 \)

- \( a = 1, b = 2, f(n) = 1 \), \( n^{\log_b a} = n^{\log_2 1} = n^0 = 1 \)
- compare \( f(n) = 1 \) with \( 1 \)
- \( 1 = \Theta(n^0) \) (so \( f(n) \) is polynomially equal to \( n^{\log_b a} \))
- case 2 applies: \( T(n) \in \Theta(n \log n) \)
### Solving Recurrences: Alt. Master Method

**Example 1a:** \( T(n) = 9T(n/3) + n \)
- \( a = 9, b = 3, k = 1, p = 0, \log_3 9 = 2 \)
- compare \( a = 9 \) with \( b^k = 3^1 = 3 \)
  \( 9 > 3 \)
- Case 1 applies: \( T(n) \in \Theta(n^{\log_3 9}) \in \Theta(n^2) \)

**Example 2a:** \( T(n) = T(n/2) + 1 \)
- \( a = 1, b = 2, k = 0, p = 0, \log_2 1 = 0 \)
- compare \( a = 1 \) with \( b^k = 2^0 = 1 \)
  \( a = b^k \) because \( 1 = 1 \)
- Case 2(a) applies: \( T(n) \in \Theta(n^{\log_2 1}) = (n^{0+\epsilon}n) = (\lg n) \in \Theta(\lg n) \)

**Example:** \( T(n) = T(n/2) + n \)
- \( a = 1, b = 2, k = 2, p = 0, \) and \( n^{\log_2 1} = n^0 = 1 \)
- compare \( a = 1 \) with \( b^k = 2^2 = 4 \)
  \( a < b^k \)
- Case 3) holds and \( T(n) = \Theta(n^{\log_2 1}) = \Theta(n^0) \)

**Example:** \( T(n) = 4T(n/2) + n \)
- \( a = 4, b = 2, k = 2, p = 0, \) and \( n^{\log_2 4} = n^2 \)
- compare \( a = 4 \) with \( b^k = 2^2 = 4 \)
  \( a = b^k \) because \( 4 = 4 \)
- Case 1 holds and \( T(n) \in \Theta(n^{\log_2 4}) \)

### Solving Recurrences: Master Method

**Example:** \( T(n) = T(n/2) + n^2 \)
- \( a = 1, b = 2, f(n) = n^2, n^{\log_2 b} = n^1 = n \)
- compare \( f(n) = n^2 \) with \( n = \Omega(n) \) (so \( f(n) \) is polynomially larger)
  \( f(n) \) is a polynomial in \( n \), case 3 holds, \( T(n) \in \Theta(n^2) \)

**Example:** \( T(n) = 7T(n/3) + n \)
- \( a = 7, b = 3, k = 2, p = 0, \) and \( n^{\log_3 7} = n^{1+\epsilon} \)
- compare \( a = 7 \) with \( b^k = 3^2 = 9 \)
  \( a > b \) because \( 7 > 4 \)
- Case 1 holds and \( T(n) \in \Theta(n^{\log_3 7}) \)

**Example:** \( T(n) = 4T(n/2) + n \)
- \( a = 4, b = 2, f(n) = n^2, n^{\log_2 b} = n^1 = n \)
- compare \( f(n) = n^2 \) with \( n = \Omega(n) \) (so \( f(n) \) is polynomially equal)
  \( f(n) \) is a polynomial in \( n \), case 2 holds and \( T(n) \in \Theta(n^2) \)

**Example:** \( T(n) = 7T(n/3) + n \)
- \( a = 7, b = 3, k = 2, p = 0, \) and \( n^{\log_3 7} = n^{1+\epsilon} \)
- compare \( a = 7 \) with \( b^k = 3^2 = 9 \)
  \( a < b^k \)
- Case 3 holds and \( T(n) \in \Theta(n^{\log_3 7}) \)