Probabilistic Recurrence Relations for Divide-and-Conquer Algorithms

Wolf Zimmermann
Brandeis University
September 21, 1992
[summary by Paul Zimmermann]

Abstract

Probabilistic recurrence relations occur frequently in the analysis of randomized algorithms. After an introduction to the work of Karp on probabilistic recurrence relations of the form

\[ T(n) = a(n) + T(h(n)) \]

where \( h(n) \) is a random variable, we discuss probabilistic recurrence relations for divide-and-conquer algorithms. In (sequential and parallel) divide-and-conquer algorithms a problem of size \( n \) is usually divided in subproblems of size \( h_1(n), \ldots, h_k(n) \) where the \( h_i(n) \) are (not necessarily independent) random variables. The analysis of the time complexity of parallel divide-and-conquer algorithms or the space complexity of sequential divide-and-conquer algorithms leads to a recurrence of the form

\[ T(n) = a(n) + \max(T(h_1(n)), \ldots, T(h_k(n))). \]

On the other hand, the analysis of the time complexity of sequential divide-and-conquer algorithms and the space complexity of parallel divide-and-conquer algorithms leads to a recurrence of the form

\[ T(n) = a(n) + T(h_1(n)) + \cdots + T(h_k(n)). \]

For both types of probabilistic recurrences, we give an upper bound for the probability distribution on \( a(n) \), if the distribution of the \( h_i(n) \) is unknown. The only informations needed are upper bounds on the expected values of the \( h_i(n) \). (Joint work with Marek Karpinski.)

The problem here is to find an upper bound for \( T(n) \), with little information about the distribution of \( h(n), h_1(n), \ldots, h_k(n) \).

Karp studied Equation (1) with the following assumptions [1]:

(i) \( a(n) \geq 0 \),
(ii) \( h(n) \) is a random variable over \([0, n]\),
(iii) \( E[h(n)] \leq m(n) \) where \( 0 \leq m(n) < n \), and \( m(n) \) and \( m(n)/n \) are non-decreasing.

Under these conditions, if one defines \( u(n) \) to be the least non-negative solution of

\[ \tau(n) = a(n) + \tau(m(n)), \]

we have two results according to the function \( a(n) \).
IV Analysis of Algorithms and Data Structures

Theorem 1. If $a(n) = 0$ for $n < d$ and $a(n) = 1$ for $n \geq d$, let $c_t = \min\{x \mid u(x) \geq t\}$, then

$$\Pr[T(n) \geq u(n) + n] \leq \left(\frac{m(n)}{n}\right)^{n-1} \frac{m(n)}{c_{u(n)}}.$$

Theorem 2. If $a(n)$ is continuous and strictly increasing, then $m(n)$ is continuous and

$$\Pr[T(n) \geq u(n) + na(n)] \leq \left(\frac{m(n)}{n}\right)^n.$$

In the general case of equation 2, Wolf Zimmermann and Marek Karpinski obtained the following result [2].

Theorem 3. Let $a(n)$ be continuous, non-decreasing, and strictly increasing on $\{n \mid a(n) > 0\}$. Also let the $m_i(n)$ be strictly increasing. Then, for every instance $z$ of size $n$ and every positive integer $j$, we have

$$\Pr[T(z) \geq u(n) + ja(n)] \leq \left(\frac{m_1(n) + \cdots + m_k(n)}{n}\right)^j.$$

Bibliography
