Skiplists

Burton Rosenberg

November 15, 2004

Skiplists are a randomized data structure supporting logarithmic-time searching. It is a linked list of nodes, each node included in layers zero up to k, where k is chosen randomly with a geometric distribution when the node is inserted. In practice, a coin is flipped and the node is inserted into the next layer up for as long as the coin flip yields heads. With a coin of bias p, this means that about a fraction p of nodes are filtered from layer k and linked together in layer $k+1$. The unchosen nodes are "skipped over".

Typically p is chosen to be a constant, and the result is $\log n$ layers for a skiplist of n nodes, both in expectation and with high probability. This is proven below. Given that a skiplist as k layers, the search time is calculated by considering the backwards process of "climbing out" of a skiplist from the found node. Each backward step is either to a node on the same level or upwards a level, with a Bernoulli distribution of probability p. Hence the number of nodes is distributed as the number of trials before the k-th success. This is distributed according to the negative binomial distribution, giving a search time of $\Theta(\log n)$.

Suppose X_i are i.i.d. geometric random variables and $Y_n = \max(X_1, \ldots, X_n)$. By independence,

$$P(Y_n \leq k) = P(X_i \leq k,\ i = 1, \ldots, n) = (P(X_1 \leq k))^n = (1 - q^k)^n.$$

and,

$$E(Y_n) = \sum_{k \geq 0} P(Y_n > k) = \sum_{k \geq 0} 1 - (1 - q^k)^n.$$

Because $(1-x)^n$ is convex upwards, and tangent to the line $1 - nx$ at zero, $(1-x)^n \geq 1 - nx$. If we set,

$$m = \frac{\log n}{\log(1/q)},$$

then $q^m = 1/n$, and

$$1 - (1 - q^{m+j})^n = 1 - (1 - q^j/n)^n \leq 1 - (1 - q^j) = q^j$$

With high probability, Y_n will be αm or less, for all $\alpha > 1$,

$$P(Y_n > \alpha m) = 1 - (1 - q^{m+(\alpha-1)m})^n \leq q^{(\alpha-1)m} = 1/n^{\alpha-1}$$

We can therefore break up the sum,

$$\begin{aligned}
\sum_{k\geq 0} 1-(1-q^k)^n &= \sum_{k=0}^{m} 1-(1-q^k)^n + \sum_{j\geq 1} 1-(1-q^{m+j})^n \\
&\leq \sum_{k=0}^{m} 1 + \sum_{j\geq 1} q^j \\
&= m+1/p
\end{aligned}$$