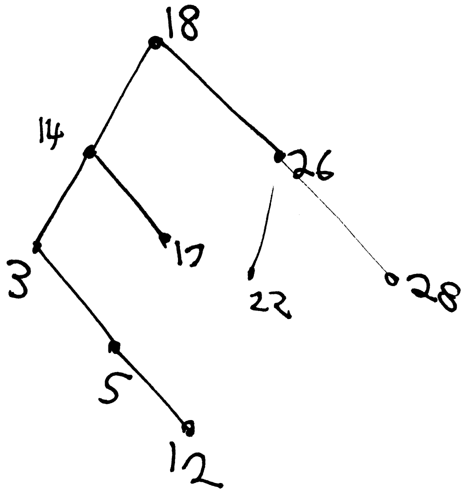
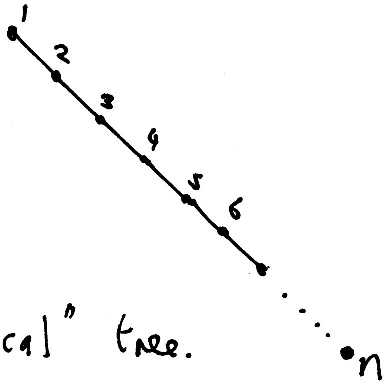
09/29/06

Random Binary Search Trees

18, 14, 17, 3, 5, 26, 12, 28, 22



"Efficiency" is "proportional" to the height.

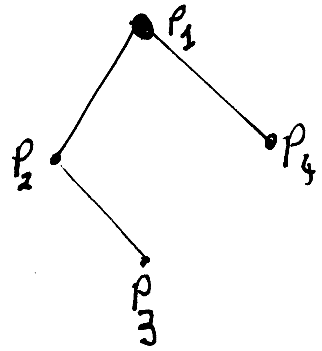


In the worst case you can get depth n after adding n items

How about a "typical" tree.

Define typical.

Tree is determined by the input sequence.



Put in n "particles".

??TRIE??

What is the likely depth?

We show that with probability $\to 1$

as $n \to \infty$ the depth is $O(\log n)$.

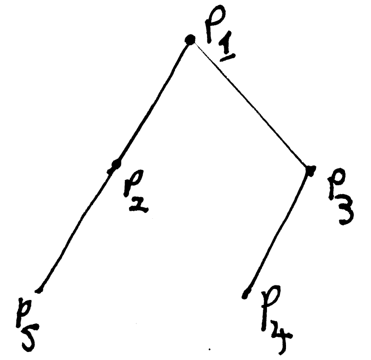
D_n is the depth of tree after adding n points.

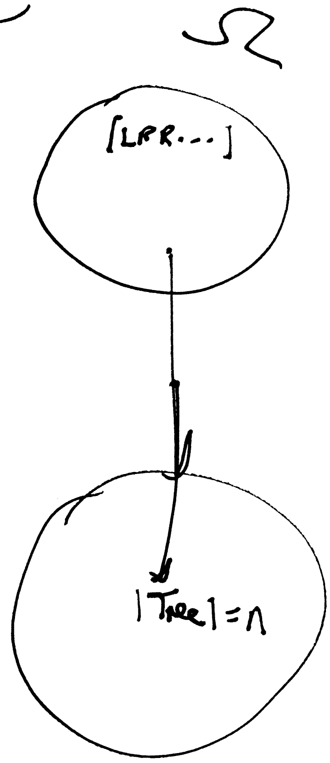
$$\Pr(D_n \geq t) \leq \left(n\, 2^{-(t-1)/2}\right)^t$$

So if $t \gg 2\log_2 n$, this is small.

$$\Omega = \{\underbrace{LRRLLLLR.\ \ldots}_{n^2}$$

$$|\Omega| = 2^{n^2}$$





(1) I don't need all bits. ~~usually~~

$\mathrm{DEEP} = \{ D_n \geq t \}$

$\Pr(\mathrm{DEEP}) \leq$??...

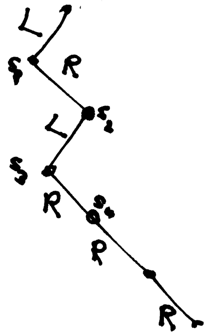
So I use Boole's Inequality (Union Bound)

$$\mathrm{DEEP} = \bigcup_{\text{PATHS } P} \; \bigcup_{\substack{S \subseteq [n] \\ |S| = t}} \underbrace{\mathrm{DEEP}(P, S)}$$

{ particles $s_1, s_2, \dots s_t \in S$ build P

$|P| = t$

LRLRRR



$$P[\mathrm{DEEP}] \leq \sum_{p} \sum_{S} P[\mathrm{DEEP}(p,S)]$$

$$= \sum_{P} \sum_{S} 2^{-[1+2+3+\cdots+t]}$$

\# choices $= 2^t$

$$= \sum_{P} \sum_{S} 2^{-t(t+1)/2}$$

\#choices $\binom{n}{t}$

$$= 2^t \binom{n}{t} 2^{-t(t+1)/2}$$

$$\leq 2^t n^t 2^{-t(t+1)/2}$$

$\binom{n}{t} \leq n^t$

$$= \left(n\, 2^{-(t-1)/2}\right)^t.$$