

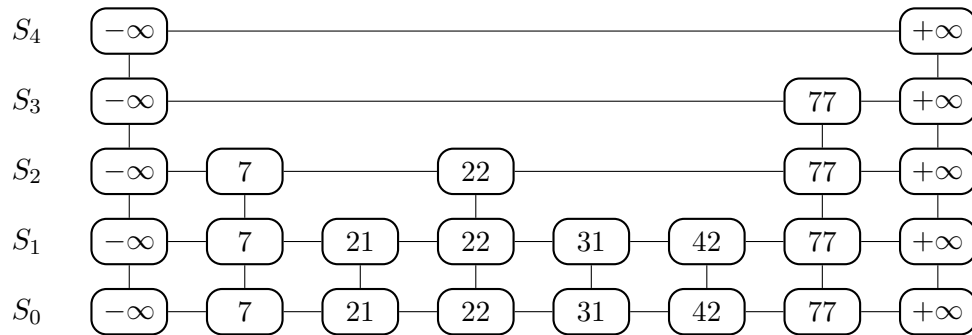
## Assignment 3 Problem 4

- a) Practice (not worth any marks): Starting with an empty skip list, insert the following keys in order: 22 31 77 7 21 42 using the following flip sequence:

HHTHTHHHTHHTHTHT

You should obtain the skip list given in the next part.

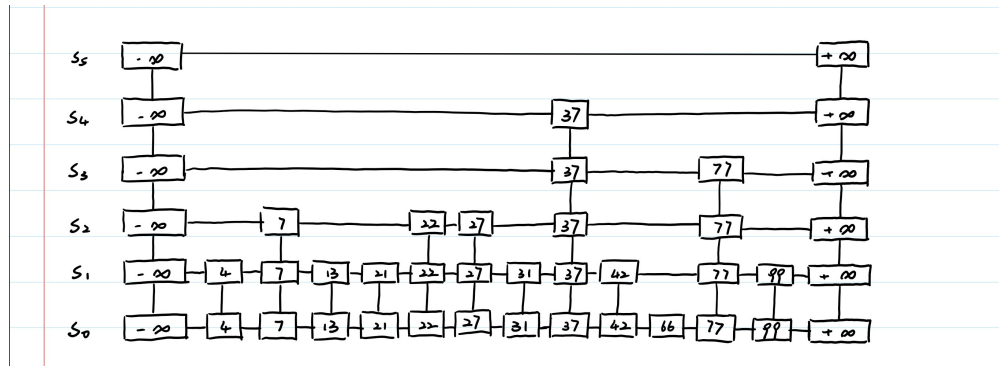
- b) Given the following skip list:



Show the skip list that is created by inserting the keys: 37, 66, 13, 4, 27, 99 into the given skip list. You must use the following coin flip sequence:

HHHHTTHTHTHHTHTTTHHHHHTHHTT

Note that each coin flip in the sequence will only be used once, in order and there may be some unused coin flips.



After inserting all keys, determine the exact number of comparisons (between 2 keys) required to search for the keys inserted in this part (that is, indicate what the search cost would be for 37, and then the search cost for 66, and so on).

For example, a search for 18 in the given skip list above, requires 6 comparisons.

Key	37	66	13	4	27	99
Comparisons	10	8	7	6	8	8

For the remaining parts, assume that the probability of adding a level to a tower is  $p$  ( $0 < p < 1$ ), as opposed to  $\frac{1}{2}$ .

- c) Explain why the probability that a key in the skip list has height at least  $i$  is  $p^i$ .

Since the probability of adding a level to a tower is  $p$ , therefore, when we need a key that appears in the level  $i$ , the probability is  $\underbrace{p \times p \times \dots \times p}_{i \text{ } p's} = p^i$ , which also means

that the probability of this key has height at least  $i$  should be  $p^i$  (since it appears in level  $i$ )

- d) Show that the expected number of extra nodes (i.e., the total number of nodes in skip list not including the nodes in  $S_0$ ) is  $O(n)$ . Therefore, the space requirements for this skip list are linear in the number of keys being stored.

Let  $|S_i|$  be the number of keys in list  $S_i$

Let  $X_k$  be the height of tower of key  $k$ , and  $P(X_k \geq i) = p^i$

Let  $l_{i,k} = 0$  if  $X_k < i$  and  $l_{i,k} = 1$  if  $X_k \geq i$

$$|S_i| = \sum_{\text{key } k} (l_{i,k})$$

$$E(|S_i|) = E(\sum_{\text{key } k} (l_{i,k})) = \sum_{\text{key } k} E(l_{i,k}) = \sum_{\text{key } k} P(l_{i,k} = 1) = \sum_{\text{key } k} P(X_k \geq i) = np^i$$

Then, expected space for extra nodes should be  $E(\sum_{i \geq 1} |S_i|) = \sum_{i \geq 1} np^i = \frac{np}{1-p} = \frac{p}{1-p} \times n \in O(n)$