Dynamic Search Structures

Self-balancing Trees and Skip Lists

Benjamin Sach
Dynamic Search Structures

A dynamic search structure, stores a set of elements

Each element $x$ must have a unique key - $x.key$

The following operations are supported:

$\text{INSERT}(x, k)$ - inserts $x$ with key $k = x.key$

$\text{FIND}(k)$ - returns the (unique) element $x$ with $x.key = k$

(or reports that it doesn’t exist)

$\text{DELETE}(k)$ - deletes the (unique) element $x$ with $x.key = k$

(or reports that it doesn’t exist)
Dynamic Search Structures

A dynamic search structure, stores a set of elements

Each element \( x \) must have a unique key - \( x.key \)

The following operations are supported:

\[
\text{INSERT}(x, k) \quad \text{inserts } x \text{ with key } k = x.key
\]

\[
\text{FIND}(k) \quad \text{returns the (unique) element } x \text{ with } x.key = k
\]

\[
\text{DELETE}(k) \quad \text{deletes the (unique) element } x \text{ with } x.key = k
\]

We would also like it to support (among others):

\[
\text{PREDECESSOR}(k) \quad \text{returns the (unique) element } x
\]

\[
\text{RANGEFIND}(k_1, k_2) \quad \text{returns every element } x \\text{ with } k_1 \leq x.key \leq k_2
\]
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren’t all efficient

Let $n$ denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as $n$ grows
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... \textit{but they aren't all efficient}

Let $n$ denote the number of elements stored in the structure
- \textit{our goal is to implement a structure with operations which scale well as $n$ grows}

We could implement a Dynamic Search Structure using an unsorted linked list:
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... 

but they aren't all efficient

Let $n$ denote the number of elements stored in the structure.

- our goal is to implement a structure with operations which scale well as $n$ grows

We could implement a Dynamic Search Structure using an unsorted linked list:
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure…

*but they aren’t all efficient*

Let \( n \) denote the number of elements stored in the structure

- *our goal is to implement a structure with operations which scale well as \( n \) grows*

We could implement a Dynamic Search Structure using an unsorted linked list:

\[
\text{Emma} \rightarrow \text{Bob} \rightarrow \text{Dawn} \rightarrow \text{Alice}
\]

**INSERT** is very efficient,

- add the new item to the head of the list in \( O(1) \) time
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure…

but they aren’t all efficient

Let \( n \) denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as \( n \) grows

We could implement a Dynamic Search Structure using an unsorted linked list:

\[
\begin{array}{cccccc}
\text{Chris} & \text{Emma} & \text{Bob} & \text{Dawn} & \text{Alice} \\
7 & 6 & 5 & 4 & 3
\end{array}
\]

**INSERT** is very efficient,

- add the new item to the head of the list in \( O(1) \) time
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren't all efficient

Let \( n \) denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as \( n \) grows

We could implement a Dynamic Search Structure using an unsorted linked list:

\[
\begin{array}{c}
\text{Chris} & \rightarrow & \text{Emma} & \rightarrow & \text{Bob} & \rightarrow & \text{Dawn} & \rightarrow & \text{Alice} \\
7 & \rightarrow & 6 & \rightarrow & 5 & \rightarrow & 4 & \rightarrow & 3 \\
\end{array}
\]

**INSERT** is very efficient,

- add the new item to the head of the list in \( O(1) \) time

**FIND** and **DELETE** are very inefficient, they take \( O(n) \) time

- we have to look through the entire linked list to find an item (in the worst case)
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... 

but they aren’t all efficient

Let \( n \) denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as \( n \) grows

We could implement a Dynamic Search Structure using an unsorted linked list:

\[
\begin{array}{cccccc}
\text{Chris} & \rightarrow & \text{Emma} & \rightarrow & \text{Bob} & \rightarrow & \text{Dawn} & \rightarrow & \text{Alice} \\
7 & \rightarrow & 6 & \rightarrow & 5 & \rightarrow & 4 & \rightarrow & 3 \\
\end{array}
\]

**INSERT** is very efficient,

- add the new item to the head of the list in \( O(1) \) time

**FIND** and **DELETE** are very inefficient, they take \( O(n) \) time

- we have to look through the entire linked list to find an item (in the worst case)
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren't all efficient

Let $n$ denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as $n$ grows

We could implement a Dynamic Search Structure using an unsorted linked list:

- `INSERT` is very efficient,
  - add the new item to the head of the list in $O(1)$ time

- `FIND` and `DELETE` are very inefficient, they take $O(n)$ time
  - we have to look through the entire linked list to find an item (in the worst case)
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... **but they aren’t all efficient**

Let $n$ denote the number of elements stored in the structure
- *our goal is to implement a structure with operations which scale well as $n$ grows*

We could implement a Dynamic Search Structure using an unsorted linked list:

![Diagram of linked list with nodes labeled Chris, Emma, Bob, Dawn, Alice, and a box labeled $x$ with $x.key$]

**INSERT** is very efficient,
- add the new item to the head of the list in $O(1)$ time

**FIND** and **DELETE** are very *inefficient*, they take $O(n)$ time
- we have to look through the entire linked list to find an item *(in the worst case)*
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren't all efficient

Let \( n \) denote the number of elements stored in the structure
- *our goal is to implement a structure with operations which scale well as \( n \) grows*

We could implement a Dynamic Search Structure using an unsorted linked list:

\[
\begin{array}{ccccc}
\text{Chris} & \rightarrow & \text{Emma} & \rightarrow & \text{Bob} & \rightarrow & \text{Dawn} & \rightarrow & \text{Alice} \\
7 & \rightarrow & 6 & \rightarrow & 5 & \rightarrow & 4 & \rightarrow & 3 \\
\end{array}
\]

**INSERT** is very efficient,
- add the new item to the head of the list in \( O(1) \) time

**FIND** and **DELETE** are very *inefficient*, they take \( O(n) \) time
- we have to look through the entire linked list to find an item *(in the worst case)*
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren’t all efficient

Let $n$ denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as $n$ grows

We could implement a Dynamic Search Structure using an unsorted linked list:

![Diagram of a linked list with nodes Chris (7), Emma (6), Bob (5), Dawn (4), Alice (3).]

**INSERT** is very efficient,

- add the new item to the head of the list in $O(1)$ time

**FIND** and **DELETE** are very inefficient, they take $O(n)$ time

- we have to look through the entire linked list to find an item (in the worst case)
Using a Linked List as a Dynamic Search Structure

There are many ways in which we could implement a search structure... but they aren’t all efficient

Let $n$ denote the number of elements stored in the structure

- our goal is to implement a structure with operations which scale well as $n$ grows

We could implement a Dynamic Search Structure using an unsorted linked list:

INSERT is very efficient,

- add the new item to the head of the list in $O(1)$ time

FIND and DELETE are very inefficient, they take $O(n)$ time

- we have to look through the entire linked list to find an item (in the worst case)
A classic choice for a dynamic search structure is a binary search tree...
A classic choice for a dynamic search structure is a binary search tree…

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
A classic choice for a dynamic search structure is a binary search tree. In a binary search tree:

- All the nodes in the left subtree have smaller keys.
- All the nodes in the right subtree have larger keys.

Recall that in a binary search tree, for any node:

- All the nodes in the left subtree have smaller keys.
- All the nodes in the right subtree have larger keys.
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
A classic choice for a dynamic search structure is a binary search tree…

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree... 

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a `FIND` operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a $\text{FIND}$ operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

21 < 27 so go left
21 > 20 so go right

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node:

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
A classic choice for a dynamic search structure is a binary search tree...

![Binary Search Tree Diagram]

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height.
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
this takes $O(h)$ time - where $h$ is the height

The INSERT and DELETE operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root...

this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

21 < 27 so go left

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The INSERT and DELETE operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

21 < 27 so go left

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...
this takes $O(h)$ time - where $h$ is the height

The INSERT and DELETE operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes \( O(h) \) time - where \( h \) is the height

The **INSERT** and **DELETE** operations are similar and also take \( O(h) \) time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a \textbf{FIND} operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The \textbf{INSERT} and \textbf{DELETE} operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree.

- 21 < 27 so go left
- 21 > 20 so go right
- 21 < 23 so go left

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root...

this takes $O(h)$ time - where $h$ is the height

The INSERT and DELETE operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **Find** operation by following a path from the root...

this takes $O(h)$ time - where $h$ is the height

The **Insert** and **Delete** operations are similar and also take $O(h)$ time.
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time.
A classic choice for a dynamic search structure is a binary search tree…

21 < 27 so go left
21 > 20 so go right
21 < 23 so go left
21 = 21 found it!
now delete it!

Recall that in a binary search tree, for any node
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root…
this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node:

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time.
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a FIND operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The INSERT and DELETE operations are similar and also take $O(h)$ time
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node

- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree...

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
A classic choice for a dynamic search structure is a binary search tree.

Recall that in a binary search tree, for any node:
- all the nodes in the left subtree have smaller keys
- all the nodes in the right subtree have larger keys

We perform a **FIND** operation by following a path from the root... this takes $O(h)$ time - where $h$ is the height

The **INSERT** and **DELETE** operations are similar and also take $O(h)$ time
Binary Search Trees

A classic choice for a dynamic search structure is a binary search tree…
A classic choice for a dynamic search structure is a binary search tree…

How big is $h$?
A classic choice for a dynamic search structure is a binary search tree…

$n$ is the number of elements

How big is $h$?
A classic choice for a dynamic search structure is a binary search tree...

How big is $h$?

It might be as small as $\log_2 n$ (if the tree is perfectly balanced)
A classic choice for a dynamic search structure is a binary search tree...

\( n \) is the number of elements

How big is \( h \)?

It might be as small as \( \log_2 n \) (if the tree is perfectly balanced)

but it might be as big as \( n \) (if the tree is completely unbalanced)
A classic choice for a dynamic search structure is a binary search tree…

$h$ is the number of elements

How big is $h$?

It might be as small as $\log_2 n$ (if the tree is perfectly balanced)

but it might be as big as $n$ (if the tree is completely unbalanced)

In particular, each INSERT could increase $h$ by one
A classic choice for a dynamic search structure is a binary search tree...

\( n \) is the number of elements

How big is \( h \)?

It might be as small as \( \log_2 n \) \textit{(if the tree is perfectly balanced)}

but it might be as big as \( n \) \textit{(if the tree is completely unbalanced)}

In particular, each \texttt{INSERT} could increase \( h \) by one
A classic choice for a dynamic search structure is a binary search tree…

$n$ is the number of elements

How big is $h$?

It might be as small as $\log_2 n$ (if the tree is perfectly balanced)
but it might be as big as $n$ (if the tree is completely unbalanced)

In particular, each INSERT could increase $h$ by one
A classic choice for a dynamic search structure is a binary search tree…

\( n \) is the number of elements

How big is \( h \)?

It might be as small as \( \log_2 n \) (if the tree is perfectly balanced)

but it might be as big as \( n \) (if the tree is completely unbalanced)

In particular, each \text{INSERT} could increase \( h \) by one
A classic choice for a dynamic search structure is a binary search tree...

$n$ is the number of elements

How big is $h$?

It might be as small as $\log_2 n$ (if the tree is perfectly balanced)

but it might be as big as $n$ (if the tree is completely unbalanced)

In particular, each INSERT could increase $h$ by one

how can we overcome this?
Part one
Self-balancing trees

inspired by slides by Inge Li Gørtz
in turn inspired by slides by Kevin Wayne
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children \((\text{hence the name})\)

**Perfect balance** - every path from the root to a leaf has the same length
\((\text{always, all the time})\)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children \((hence\ the\ name)\)

**Perfect balance** - every path from the root to a leaf has the same length \((always,\ all\ the\ time)\)

- **2-node:** 2 children and 1 key
- **3-node:** 3 children and 2 keys
- **4-node:** 4 children and 3 keys
**2-3-4 Trees**

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

2-node: 2 children and 1 key

3-node: 3 children and 2 keys

4-node: 4 children and 3 keys
Key idea: Nodes can have between 2 and 4 children (hence the name)

Perfect balance - every path from the root to a leaf has the same length (always, all the time)

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

| 2-node: | 2 children and 1 key |
| 3-node: | 3 children and 2 keys |
| 4-node: | 4 children and 3 keys |
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

<table>
<thead>
<tr>
<th>2-node</th>
<th>3-node</th>
<th>4-node</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 children and 1 key</td>
<td>3 children and 2 keys</td>
<td>4 children and 3 keys</td>
</tr>
</tbody>
</table>

The ● are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

Key idea: Nodes can have between 2 and 4 children (hence the name)

Perfect balance - every path from the root to a leaf has the same length (always, all the time)

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The ● are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

---

**2-node:** 2 children and 1 key  
**3-node:** 3 children and 2 keys  
**4-node:** 4 children and 3 keys

Like in a binary search tree,  
the keys held at a node determine  
the contents of its subtrees

The ◦ are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The ● are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea**: Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

---

**2-node**: 2 children and 1 key

**3-node**: 3 children and 2 keys

**4-node**: 4 children and 3 keys

*Like in a binary search tree, the keys held at a node determine the contents of its subtrees*

*The ● are “dummy leaves” (they don’t do or contain anything)*
2-3-4 Trees

Key idea: Nodes can have between 2 and 4 children (hence the name)

Perfect balance - every path from the root to a leaf has the same length (always, all the time)

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys

The ● are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

---

**2-node:** 2 children and 1 key

**3-node:** 3 children and 2 keys

**4-node:** 4 children and 3 keys

*Like in a binary search tree, the keys held at a node determine the contents of its subtrees*

*The ● are “dummy leaves” (they don’t do or contain anything)*
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

---

**2-node:** 2 children and 1 key

**3-node:** 3 children and 2 keys

**4-node:** 4 children and 3 keys

*Like in a binary search tree, the keys held at a node determine the contents of its subtrees*

*The ○ are “dummy leaves” (they don’t do or contain anything)*
**2-3-4 Trees**

**Key idea:** Nodes can have between **2** and **4** children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

---

**2-node:** 2 children and 1 key

**3-node:** 3 children and 2 keys

**4-node:** 4 children and 3 keys

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The ● are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

---

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The ○ are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

---

**2-node:** 2 children and 1 key  
**3-node:** 3 children and 2 keys  
**4-node:** 4 children and 3 keys

---

Like in a binary search tree,  
the keys held at a node determine  
the contents of its subtrees

The ● are “dummy leaves” *(they don’t do or contain anything)*
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children \(^{(hence\ the\ name)}\)

**Perfect balance** - every path from the root to a leaf has the same length \(^{(always,\ all\ the\ time)}\)

*2-node:* 2 children and 1 key  
*3-node:* 3 children and 2 keys  
*4-node:* 4 children and 3 keys

*Like in a binary search tree,*  
*the keys held at a node determine*  
*the contents of its subtrees*

*The \(\bullet\) are “dummy leaves”* (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

2-node: 2 children and 1 key
3-node: 3 children and 2 keys
4-node: 4 children and 3 keys

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The dots are “dummy leaves” (they don’t do or contain anything)
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children *(hence the name)*

**Perfect balance** - every path from the root to a leaf has the same length *(always, all the time)*

---

**Nodes:**
- **2-node:** 2 children and 1 key
- **3-node:** 3 children and 2 keys
- **4-node:** 4 children and 3 keys

---

*Like in a binary search tree, the keys held at a node determine the contents of its subtrees*

*The ovals are “dummy leaves” (they don’t do or contain anything)*
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children (hence the name)

**Perfect balance** - every path from the root to a leaf has the same length (always, all the time)

Like in a binary search tree, the keys held at a node determine the contents of its subtrees.
2-3-4 Trees

**Key idea:** Nodes can have between 2 and 4 children \((hence\ the\ name)\)

**Perfect balance** - every path from the root to a leaf has the same length \((always,\ all\ the\ time)\)

<table>
<thead>
<tr>
<th>Node Type</th>
<th>Children</th>
<th>Keys</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-node</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3-node</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4-node</td>
<td>4</td>
<td>3</td>
</tr>
</tbody>
</table>

Like in a binary search tree, the keys held at a node determine the contents of its subtrees

The dots are “dummy leaves” (they don’t do or contain anything)
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root. . .

decisions are made by inspecting the key(s) at the current node and following the appropriate edge
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

![Decision tree diagram]

Decisions are made by inspecting the key(s) at the current node and following the appropriate edge.
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

decisions are made by inspecting the key(s) at the current node and following the appropriate edge.
Just like in a binary search tree, we perform a **find** operation by following a path from the root.

12 is between 11 and 18

*decisions are made by inspecting the key(s) at the current node and following the appropriate edge*
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18

*decisions are made by inspecting the key(s) at the current node and following the appropriate edge*
The \textbf{FIND} operation

Just like in a binary search tree, we perform a \textbf{FIND} operation by following a path from the root. . .

\textbf{FIND}(12)

decisions are made by inspecting the key(s) at the current node and following the appropriate edge
The FIND operation

Just like in a binary search tree, we perform a FIND operation by following a path from the root...

12 is between 11 and 18
12 is smaller than 13

decisions are made by inspecting the key(s) at the current node and following the appropriate edge
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18  
12 is smaller than 13

*decisions are made by inspecting the key(s) at the current node and following the appropriate edge*
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18  
12 is smaller than 13  

*decisions are made by inspecting the key(s) at the current node and following the appropriate edge*
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18
12 is smaller than 13
found it!

decisions are made by inspecting the key(s) at the current node and following the appropriate edge
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root.

12 is between 11 and 18
12 is smaller than 13
found it!

**FIND**(12)

*decisions are made by inspecting the key(s) at the current node and following the appropriate edge*

*What is the time complexity of the **FIND** operation?*
The **FIND** operation

Just like in a binary search tree,
   we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18
12 is smaller than 13
found it!

**FIND**(12)

---

decisions are made by inspecting the key(s) at the current node and following the appropriate edge

*What is the time complexity of the **FIND** operation?*

It’s $O(h)$ again
The **FIND** operation

Just like in a binary search tree, we perform a **FIND** operation by following a path from the root...

12 is between 11 and 18
12 is smaller than 13
found it!

```
descisions are made by inspecting the key(s) at the current node
and following the appropriate edge
```

What is the time complexity of the **FIND** operation?

It’s $O(h)$ again

(each step down the path takes $O(1)$ time)
The FIND operation

Just like in a binary search tree,
we perform a FIND operation by following a path from the root...

12 is between 11 and 18
12 is smaller than 13
found it!

decisions are made by inspecting the key(s) at the current node
and following the appropriate edge

What is the time complexity of the FIND operation?
It’s $O(h)$ again
(each step down the path takes $O(1)$ time)

What is the height, $h$ of a 2-3-4 tree?
The height of a 2-3-4 tree

**Perfect balance** - every path from the root to a leaf has the same length

(we’ll justify this as we go along)

This implies that the height, $h$ of a 2-3-4 tree with $n$ nodes is

$$\text{Best case: } \log_4 n = \frac{\log_2 n}{2} \text{ (all 4-nodes)}$$

$$\text{Worst case: } \log_2 n \text{ (all 2-nodes)}$$

$h$ is between 10 and 20 for a million nodes
The height of a 2-3-4 tree

**Perfect balance** - every path from the root to a leaf has the same length

(we’ll justify this as we go along)

This implies that the height, \( h \) of a 2-3-4 tree with \( n \) nodes is

Best case: \( \log_4 n = \frac{\log_2 n}{2} \) (all 4-nodes)

Worst case: \( \log_2 n \) (all 2-nodes)

\( h \) is between 10 and 20 for a million nodes

The time complexity of the FIND operation is \( O(h) \)
The height of a 2-3-4 tree

**Perfect balance** - every path from the root to a leaf has the same length

(we’ll justify this as we go along)

This implies that the height, $h$ of a 2-3-4 tree with $n$ nodes is

Best case: $\log_4 n = \frac{\log_2 n}{2}$ (all 4-nodes)

Worst case: $\log_2 n$ (all 2-nodes)

$h$ is between 10 and 20 for a million nodes

The time complexity of the FIND operation is $O(h) = O(\log n)$
To perform \textbf{INSERT}(x, k),
The **INSERT** operation

To perform **INSERT**(x, k),

**Step 1:** Search for the key k as if performing **FIND**(k).
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$. 

The **INSERT** operation

$\text{INSERT}(x, 16)$
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1**: Search for the key \(k\) as if performing **FIND**(\(k\)).
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).
The **INSERT** operation

To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$. 
To perform \texttt{INSERT}(x, k),

**Step 1:** Search for the key \( k \) as if performing \texttt{FIND}(k).

**Step 2:** If the leaf is a 2-node,
insert \((x, k)\), converting it into a 3-node.
The **INSERT** operation

To perform **INSERT**(x, k),

**Step 1:** Search for the key k as if performing **FIND**(k).

**Step 2:** If the leaf is a 2-node, insert (x, k), converting it into a 3-node
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,

insert $(x, k)$, converting it into a 3-node.
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**Step 2:** If the leaf is a 2-node,
insert \((x, k)\), converting it into a 3-node.
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**Step 2:** If the leaf is a 2-node, insert \((x, k)\), converting it into a 3-node.
To perform \( \text{INSERT}(x, k) \),

**Step 1:** Search for the key \( k \) as if performing \( \text{FIND}(k) \).

**Step 2:** If the leaf is a 2-node, insert \( (x, k) \), converting it into a 3-node.
The **INSERT** operation

To perform **INSERT**\( (x, k) \),

**Step 1**: Search for the key \( k \) as if performing **FIND**\( (k) \).

**Step 2**: If the leaf is a 2-node,
   
   insert \( (x, k) \), converting it into a 3-node.
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,

insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert $(x, k)$, converting it into a 4-node
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,

insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT**({\textit{x}, k}),

**Step 1:** Search for the key \textit{k} as if performing **FIND**({\textit{k}}).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,
   insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
   insert $(x, k)$, converting it into a 4-node
To perform \texttt{INSERT}(x, k),

\begin{itemize}
  \item \textbf{Step 1:} Search for the key \( k \) as if performing \texttt{FIND}(k).
  \item \textbf{Step 2:} If the leaf is a 2-node,
    \hspace{1cm} insert \((x, k)\), converting it into a 3-node
  \item \textbf{Step 3:} If the leaf is a 3-node,
    \hspace{1cm} insert \((x, k)\), converting it into a 4-node
\end{itemize}
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,

insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT**($x, k$),

**Step 1:** Search for the key $k$ as if performing **FIND**( $k$).

**Step 2:** If the leaf is a 2-node,

insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert $(x, k)$, converting it into a 4-node
To perform \( \text{INSERT}(x, k) \),

\textbf{Step 1:} Search for the key \( k \) as if performing \( \text{FIND}(k) \).

\textbf{Step 2:} If the leaf is a 2-node,

\quad insert \( (x, k) \), converting it into a 3-node

\textbf{Step 3:} If the leaf is a 3-node,

\quad insert \( (x, k) \), converting it into a 4-node

\textbf{Step 4:} If the leaf is a 4-node,
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1:** Search for the key \(k\) as if performing **FIND**(\(k\)).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node

**Step 4:** If the leaf is a 4-node, ????
The **INSERT** operation

To perform **INSERT**(x, k),

**Step 1:** Search for the key k as if performing **FIND**(k).

**Step 2:** If the leaf is a 2-node, insert (x, k), converting it into a 3-node.

**Step 3:** If the leaf is a 3-node, insert (x, k), converting it into a 4-node.

**Step 4:** If the leaf is a 4-node, ??? We will make sure this never happens.
We can split any 4-node into two 2-nodes if its parent isn’t a 4-node.
**Splitting 4-nodes**

We can **split** any 4-node into two 2-nodes if it’s parent isn’t a 4-node.

**Before**

**After**
**Splitting 4-nodes**

We can split any 4-node into two 2-nodes if its parent isn’t a 4-node.

**Before**

```
  41  63
  /    |
32     51
```

**After**

```
  41  63  86
  /    |
32     51
```

The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node).
We can **split** any 4-node into two 2-nodes if its parent isn’t a 4-node.

**Before**

These subtrees could have any size.

**After**

The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node.)
We can **split** any 4-node into two 2-nodes if its parent isn’t a 4-node.

**Before**

The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node).

**After**

These subtrees haven’t changed.
Splitting 4-nodes

We can split any 4-node into two 2-nodes if its parent isn’t a 4-node.

BEFORE

AFTER

The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node).

no path lengths have changed

these subtrees haven’t changed

these subtrees could have any size
**SPLITTING 4-nodes**

We can **split** any 4-node into two 2-nodes if it’s parent isn’t a 4-node.

*Before*

- 41 63
- 32
- 51
- 71 86 92

*After*

- 41 63 86
- 32
- 51
- 71
- 92

The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node).

No path lengths have changed.

*If it was perfectly balanced, it still is.*
**SPLITTING 4-nodes**

We can SPLIT any 4-node into two 2-nodes if it's parent isn't a 4-node.

![Diagram of tree before and after splitting a 4-node.](image)

**BEFORE**

*These subtrees could have any size*

**AFTER**

*The extra key is pushed up to the parent (so it won’t work if the parent is a 4-node)*

*No path lengths have changed (if it was perfectly balanced, it still is)*

**SPLIT** takes $O(1)$ time
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1:** Search for the key \(k\) as if performing **FIND**(\(k\)).

**SPLIT** 4-nodes as we go down

**Step 2:** If the leaf is a 2-node,

- insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

- insert \((x, k)\), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**(*x*, *k*),

**Step 1:** Search for the key *k* as if performing **FIND**(*k*).

**Step 2:** If the leaf is a 2-node,

insert (*x*, *k*), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert (*x*, *k*), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**($x, k$),

**Step 1:** Search for the key $k$ as if performing **FIND**($k$).

**Step 2:** If the leaf is a 2-node, insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node, insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT**(x, k),

**Step 1:** Search for the key k as if performing **FIND**(k).

**Step 2:** If the leaf is a 2-node,
    insert (x, k), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
    insert (x, k), converting it into a 4-node
To perform \textbf{INSERT}(x, k),

**Step 1:** Search for the key \( k \) as if performing \textbf{FIND}(k).

\textbf{_SPLIT} 4-nodes as we go down

**Step 2:** If the leaf is a 2-node,
insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
insert \((x, k)\), converting it into a 4-node
To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node,
insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT** \((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND** \((k)\).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node

**Note:** Split 4-nodes as we go down.

**Example:**

**Step 1:** Search for the key \(k\) as if performing **FIND** \((k)\).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node

**Diagram:**

1. **Step 1:** Search for the key \(k\).
2. **Step 2:** Insert \((x, k)\), converting the 2-node into a 3-node.
3. **Step 3:** Insert \((x, k)\), converting the 3-node into a 4-node.

**Split this!**

**Insert** \((x, 8)\)
The **INSERT** operation

**Step 1:** Search for the key $k$ as if performing **FIND**($k$).

**Step 2:** If the leaf is a 2-node, insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node, insert $(x, k)$, converting it into a 4-node

**Split** 4-nodes as we go down

To perform **INSERT**($x, k$),

**Insert**($x, 8$)

**Split** this!
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1:** Search for the key \(k\) as if performing **FIND**(\(k\)).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node
Step 1: Search for the key $k$ as if performing $\text{FIND}(k)$.

Step 2: If the leaf is a 2-node, insert $(x, k)$, converting it into a 3-node

Step 3: If the leaf is a 3-node, insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1**: Search for the key \(k\) as if performing **FIND**\((k)\).

**Step 2**: If the leaf is a 2-node,

- insert \((x, k)\), converting it into a 3-node

**Step 3**: If the leaf is a 3-node,

- insert \((x, k)\), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1:** Search for the key \(k\) as if performing **FIND**(\(k\)).

**Step 2:** If the leaf is a 2-node, insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node, insert \((x, k)\), converting it into a 4-node

**SPLIT** this!
To perform \( \text{INSERT}(x, k) \),

**Step 1:** Search for the key \( k \) as if performing \( \text{FIND}(k) \).

**Step 2:** If the leaf is a 2-node,
insert \( (x, k) \), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
insert \( (x, k) \), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**(\(x, k\)),

**Step 1:** Search for the key \(k\) as if performing **FIND**(\(k\)).

**Step 2:** If the leaf is a 2-node,

\[\text{insert } (x, k), \text{ converting it into a 3-node}\]

**Step 3:** If the leaf is a 3-node,

\[\text{insert } (x, k), \text{ converting it into a 4-node}\]
The \textbf{INSERT} operation

To perform $\text{INSERT}(x, k)$,

\textbf{Step 1:} Search for the key $k$ as if performing $\text{FIND}(k)$.
\hfill \textbf{SPLIT 4-nodes as we go down}

\textbf{Step 2:} If the leaf is a 2-node,
insert $(x, k)$, converting it into a 3-node

\textbf{Step 3:} If the leaf is a 3-node,
insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**Split** \(4\)-nodes as we go down

**Step 2:** If the leaf is a 2-node, insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node, insert \((x, k)\), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**SPLIT** 4-nodes as we go down

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node
The **INSERT** operation

To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$. **SPLIT 4-nodes as we go down**

**Step 2:** If the leaf is a 2-node,
insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
insert $(x, k)$, converting it into a 4-node
The **INSERT** operation

To perform **INSERT** \((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND** \((k)\).

**Step 2:** If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node

**OK, one more thing…**
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\(k\).

**Step 2:** If the leaf is a 2-node,
   insert \((x, k)\), converting it into a 3-node

**Step 3:** If the leaf is a 3-node,
   insert \((x, k)\), converting it into a 4-node

OK, one more thing...
To perform \textsc{INSERT}(x, k),

\textbf{Step 1:} Search for the key \( k \) as if performing \textsc{FIND}(k).

\textbf{Step 2:} If the leaf is a 2-node,

insert \((x, k)\), converting it into a 3-node

\textbf{Step 3:} If the leaf is a 3-node,

insert \((x, k)\), converting it into a 4-node

\textbf{OK, one more thing…}
The **INSERT** operation

To perform $\text{INSERT}(x, k)$,

**Step 1:** Search for the key $k$ as if performing $\text{FIND}(k)$.

**Step 2:** If the leaf is a 2-node, insert $(x, k)$, converting it into a 3-node

**Step 3:** If the leaf is a 3-node, insert $(x, k)$, converting it into a 4-node

OK, one more thing... what happens when we **SPLIT** the root?
The **INSERT** operation

To perform \( \text{INSERT}(x, k) \),

**Step 1:** Search for the key \( k \) as if performing \( \text{FIND}(k) \).

**Step 2:** If the leaf is a 2-node,

\[ \text{insert } (x, k), \text{ converting it into a 3-node} \]

**Step 3:** If the leaf is a 3-node,

\[ \text{insert } (x, k), \text{ converting it into a 4-node} \]

OK, one more thing... what happens when we SPLIT the root?
To perform \textsc{Insert}$(x, k)$,

\textbf{Step 1:} Search for the key $k$ as if performing \textsc{Find}$(k)$.

\textbf{Step 2:} If the leaf is a 2-node, insert $(x, k)$, converting it into a 3-node

\textbf{Step 3:} If the leaf is a 3-node, insert $(x, k)$, converting it into a 4-node

OK, one more thing... what happens when we \textsc{Split} the root?
The **INSERT** operation

The diagram shows a tree with nodes labeled with numbers. The operation **INSERT** is applied, resulting in the tree structure depicted.
The **INSERT** operation

**INSERT**(x, 20)

**SPLITTING** the root increases the height of the tree and increases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property:

- *i.e every path from the root to a leaf has the same length*
The **INSERT** operation

**INSERT**\((x, 20)\)

**SPLITTING** the root increases the height of the tree and increases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property

- *i.e* every path from the root to a leaf has the same length

This is the only way **INSERT** can affect the length of paths so it also maintains the **perfect balance** property.
The `INSERT` operation

**SPLITTING** the root increases the height of the tree and increases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property:

- *i.e. every path from the root to a leaf has the same length*

This is the only way `INSERT` can affect the length of paths so it also maintains the **perfect balance** property.

As each `SPLIT` takes $O(1)$ time, overall `INSERT` takes $O(\log n)$ time.
The **INSERT** operation

To perform **INSERT**\((x, k)\),

**Step 1:** Search for the key \(k\) as if performing **FIND**\((k)\).

**Step 2:** If the bottom node is a 2-node,
insert \((x, k)\), converting it into a 3-node

**Step 3:** If the bottom node is a 3-node,
insert \((x, k)\), converting it into a 4-node

As each **SPLIT** takes \(O(1)\) time, overall **INSERT** takes \(O(\log n)\) time
The **DELETE** operation

To perform **DELETE**(\(k\)) on a leaf (we’ll deal with other nodes later)
The **DELETE** operation

To perform **DELETE**(\(k\)) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).
The **DELETE** operation

**Step 1:** Search for the key \( k \) using **FIND**(\( k \)).

To perform **DELETE**\((k)\) on a leaf (we’ll deal with other nodes later)
The **DELETE** operation

To perform **DELETE**($k$) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using **FIND**($k$).
The **DELETE** operation

To perform **DELETE**($k$) on a leaf (*we’ll deal with other nodes later*)

**Step 1:** Search for the key $k$ using **FIND**($k$).
The **DELETE** operation

To perform **DELETE**($k$) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using **FIND**($k$).
To perform DELETE($k$) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using FIND($k$).

**Step 2:** If the leaf is a 3-node,

delete $(x, k)$, converting it into a 2-node
The **DELETE** operation

To perform **DELETE** \((k)\) on a leaf (*we’ll deal with other nodes later*)

**Step 1:** Search for the key \(k\) using **FIND** \((k)\).

**Step 2:** If the leaf is a 3-node,

delete \((x, k)\), converting it into a 2-node
The **DELETE** operation

To perform **DELETE**($k$) on a leaf *(we’ll deal with other nodes later)*

**Step 1:** Search for the key $k$ using **FIND**($k$).

**Step 2:** If the leaf is a 3-node,

delete $(x, k)$, converting it into a 2-node
The **DELETE** operation

To perform **DELETE**($k$) on a leaf *(we’ll deal with other nodes later)*

**Step 1:** Search for the key $k$ using **FIND($k$)**.

**Step 2:** If the leaf is a 3-node,

- delete $(x, k)$, converting it into a 2-node
The **DELETE** operation

To perform **DELETE**(\(k\)) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).

**Step 2:** If the leaf is a 3-node,
   delete \((x, k)\), converting it into a 2-node.
The **DELETE** operation

To perform **DELETE**(k) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key k using **FIND**(k).

**Step 2:** If the leaf is a 3-node, delete (x, k), converting it into a 2-node
The **DELETE** operation

To perform **DELETE**(k) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key k using **FIND**(k).

**Step 2:** If the leaf is a 3-node,

delete (x, k), converting it into a 2-node
The **DELETE** operation

To perform **DELETE**(k) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key k using **FIND**(k).

**Step 2:** If the leaf is a 3-node,
   delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

To perform **DELETE** \((k)\) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND** \((k)\).

**Step 2:** If the leaf is a 3-node, delete \((x, k)\), converting it into a 2-node.

**Step 3:** If the leaf is a 4-node, delete \((x, k)\), converting it into a 3-node.
The **DELETE** operation

To perform **DELETE**\((k)\) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND**\((k)\).

**Step 2:** If the leaf is a 3-node,

delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

---

**Step 1:** Search for the key $k$ using $\text{FIND}(k)$.

**Step 2:** If the leaf is a 3-node,
   delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete $(x, k)$, converting it into a 3-node

---

To perform **DELETE**($k$) **on a leaf** *(we’ll deal with other nodes later)*
The **DELETE** operation

To perform **DELETE**(\(k\)) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).

**Step 2:** If the leaf is a 3-node,
- delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
- delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**(\(k\)) on a leaf *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).

**Step 2:** If the leaf is a 3-node,
   delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**($k$) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key $k$ using **FIND**($k$).

**Step 2:** If the leaf is a 3-node,
delete ($x$, $k$), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
delete ($x$, $k$), converting it into a 3-node
To perform $\text{DELETE}(k)$ on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using $\text{FIND}(k)$.

**Step 2:** If the leaf is a 3-node,
   delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete $(x, k)$, converting it into a 3-node

**Step 4:** If the leaf is a 2-node,
The **DELETE** operation

To perform **DELETE**(\(k\)) **on a leaf** (we’ll deal with other nodes later)

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).

**Step 2:** If the leaf is a 3-node, 
    delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node, 
    delete \((x, k)\), converting it into a 3-node

**Step 4:** If the leaf is a 2-node,  ???
The **DELETE** operation

To perform **DELETE**($k$) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key $k$ using **FIND**($k$).

**Step 2:** If the leaf is a 3-node,
delete ($x$, $k$), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
depend ($x$, $k$), converting it into a 3-node

**Step 4:** If the leaf is a 2-node, ?? We will make sure this *never* happens
**Fusing** 2-nodes

We can **FUSE** two 2-nodes (with the same parent) into a 4-node if that parent isn’t a 2-node
**Fusing 2-nodes**

We can **Fuse** two 2-nodes (with the same parent) into a 4-node if that parent isn't a 2-node.

**Before**

```
  41 63 86
  /   /
71    92
```

**After**

```
  41 63
  /   /
71    86 92
```
**Fusing 2-nodes**

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn't a 2-node.

**BEFORE**

**AFTER**

The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node).
**Fusing 2-nodes**

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn’t a 2-node.

This is the opposite of a split operation.

**Before**

**After**

The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node).
Fusing 2-nodes

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn’t a 2-node.

This is the opposite of a split operation.

**BEFORE**

**AFTER**

The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node).

these subtrees haven’t changed
**Fusing 2-nodes**

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn’t a 2-node.

This is the opposite of a split operation.

**Before**

**After**

The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node)

no path lengths have changed

these subtrees haven’t changed
**Fusing 2-nodes**

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn’t a 2-node.

This is the opposite of a split operation.

**Before**

- The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node).
- No path lengths have changed (if it was perfectly balanced, it still is).

**After**

- These subtrees haven’t changed.
**Fusing 2-nodes**

We can fuse two 2-nodes (with the same parent) into a 4-node if that parent isn't a 2-node.

This is the opposite of a split operation.

**Before**

**After**

The extra key is pulled down from the parent (so it won’t work if the parent is a 2-node)

no path lengths have changed

(if it was perfectly balanced, it still is)

Fuse takes $O(1)$ time.
TRANSFERING keys

If there is a 2-node and a 3-node (with the same parent) we can perform a TRANSFER (even if the parent is the root)
If there is a 2-node and a 3-node (with the same parent) we can perform a transfer (even if the parent is the root)
Transfering keys

If there is a 2-node and a 3-node (with the same parent)
we can perform a Transfer (even if the parent is the root)

Before

AFTER

The keys have been rearranged
If there is a 2-node and a 3-node (with the same parent), we can perform a Transfer (even if the parent is the root).

Before

After

The keys have been rearranged.

These subtrees haven’t changed.
If there is a 2-node and a 3-node (with the same parent) we can perform a Transfer (even if the parent is the root)

The keys have been rearranged

no path lengths have changed

these subtrees haven’t changed
**TRANSFERRING** keys

If there is a 2-node and a 3-node (with the same parent) we can perform a TRANSFER (even if the parent is the root).

BEFORE

AFTER

The keys have been rearranged

no path lengths have changed

*(if it was perfectly balanced, it still is)*

these subtrees haven’t changed
TRANSFERRING keys

If there is a 2-node and a 3-node
(with the same parent)
we can perform a TRANSFER
(even if the parent is the root)

BEFORE

AFTER

The keys have been rearranged

no path lengths have changed
(if it was perfectly balanced, it still is)

TRANSFER takes $O(1)$ time
If there is a 2-node and a 3-node (with the same parent), we can perform a Transfer (even if the parent is the root).

Transfer also works with a 2-node and a 4-node...

The keys have been rearranged.

no path lengths have changed

(if it was perfectly balanced, it still is)

Transfer takes $O(1)$ time.
The **DELETE** operation

To perform **DELETE**(*k*) **on a leaf** (we’ll deal with other nodes later)

**Step 1:** Search for the key *k* using **FIND**(*k*).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete (**x**, *k*), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete (**x**, *k*), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**(\(k\)) **on a leaf** (*we’ll deal with other nodes later*)

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).
   use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
   delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete \((x, k)\), converting it into a 3-node
To perform \texttt{DELETE}(k) on a leaf (we’ll deal with other nodes later)

\textbf{Step 1:} Search for the key \textit{k} using \texttt{FIND}(k).

use \texttt{FUSE} and \texttt{TRANSFER} to convert 2-nodes as we go down

\textbf{Step 2:} If the leaf is a 3-node,

delete \((x, k)\), converting it into a 2-node

\textbf{Step 3:} If the leaf is a 4-node,

delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**(\(k\)) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND**(\(k\)).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete \((x, k)\), converting it into a 3-node
To perform **DELETE** \((k)\) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key \(k\) using **FIND** \((k)\).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

**Step 1:** Search for the key \( k \) using **FIND(\( k \))**.
- use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
- delete \( (x, k) \), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
- delete \( (x, k) \), converting it into a 3-node

To perform **DELETE(\( k \))** on a leaf *(we’ll deal with other nodes later)*

---

**Fuse this!**

---

**Delete(5)**

---

**Diagram:**

```
  11 18
  /     \
2 4 6   13 15
/           /
1 3 5       12 14
           /     \
          7 10   17
           /     \
         19 22   25 26
```

---

---
The **DELETE** operation

**Step 1:** Search for the key $k$ using **FIND**($k$).

* use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node, delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node, delete $(x, k)$, converting it into a 3-node

To perform **DELETE**($k$) on a leaf (*we’ll deal with other nodes later*)
The DELETE operation

To perform $\text{DELETE}(k)$ on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using $\text{FIND}(k)$.
use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
delete $(x, k)$, converting it into a 3-node
The **DELETE** operation

To perform **DELETE**($k$) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using **FIND**($k$).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete ($x$, $k$), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete ($x$, $k$), converting it into a 3-node
The **DELETE** operation

**Step 1:** Search for the key $k$ using **FIND**($k$).

**Step 2:** If the leaf is a 3-node,

- delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

- delete $(x, k)$, converting it into a 3-node

To perform **DELETE**($k$) on a leaf (we'll deal with other nodes later)
The **DELETE** operation

To perform **DELETE**(\(k\)) **on a leaf** (we’ll deal with other nodes later)

**Step 1:** Search for the key \(k\) using **FIND**\((k)\).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete \((x, k)\), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete \((x, k)\), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**(*k*) on a leaf (we'll deal with other nodes later)

**Step 1:** Search for the key *k* using **FIND**(k).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete (*x*, *k*), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete (*x*, *k*), converting it into a 3-node
The **DELETE** operation

To perform **DELETE**(k) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key k using **FIND**(k).

  use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

  delete (x, k), converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

  delete (x, k), converting it into a 3-node
The **DELETE** operation

Step 1: Search for the key $k$ using **FIND**($k$).

Step 2: If the leaf is a 3-node, delete $(x, k)$, converting it into a 2-node.

Step 3: If the leaf is a 4-node, delete $(x, k)$, converting it into a 3-node.

To perform **DELETE($k$)** on a leaf (**we’ll deal with other nodes later**)

Now we can **DELETE** the 5.
The DELETE operation

To perform DELETE($k$) on a leaf (we’ll deal with other nodes later)

**Step 1:** Search for the key $k$ using FIND($k$).
   use FUSE and TRANSFER to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
   delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete $(x, k)$, converting it into a 3-node

Now we can DELETE the 5
The **DELETE** operation

To perform **DELETE**($k$) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key $k$ using **FIND**($k$).
   use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
   delete $(x, k)$, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
   delete $(x, k)$, converting it into a 3-node
The \textbf{DELETE} operation

To perform \textbf{DELETE}(k) on a leaf \textit{(we'll deal with other nodes later)}

\textbf{Step 1:} Search for the key \textit{k} using \textbf{FIND}(k).  
use \textbf{FUSE} and \textbf{TRANSFER} to convert 2-nodes as we go down

\textbf{Step 2:} If the leaf is a 3-node,  
delete \textit{(x, k)}, converting it into a 2-node

\textbf{Step 3:} If the leaf is a 4-node,  
delete \textit{(x, k)}, converting it into a 3-node

\textit{OK, one more thing…}
The **DELETE** operation

To perform **DELETE**(k) **on a leaf** *(we’ll deal with other nodes later)*

**Step 1:** Search for the key *k* using **FIND**(k).

use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,

delete *(x, k)*, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,

delete *(x, k)*, converting it into a 3-node

OK, one more thing... what happens when we **FUSE** the root?
**Fusing** the root

We said that we could only **fuse** two 2-nodes if the parent was not a 2-node... we make an exception for the root.

**Fusing** the root can decrease the height of the tree which in turn decreases the length of all root-leaf paths by one.
**Fusing** the root

We said that we could only **Fuse** two 2-nodes if the parent was not a 2-node... we make an exception for the root.

![Tree diagram]

**Fusing** the root can decrease the height of the tree which in turn decreases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property:

- *i.e every path from the root to a leaf has the same length*
**FUSING** the root

We said that we could only **Fuse** two 2-nodes if the parent was not a 2-node... we make an exception for the root.

FUSING the root can decrease the height of the tree which in turn decreases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property

- *i.e every path from the root to a leaf has the same length*

This is the only way **Delete** can affect the length of paths so it also maintains the **perfect balance** property.
**Fusing** the root

We said that we could only **Fuse** two 2-nodes if the parent was not a 2-node... we make an exception for the root.

Fusing the root can decrease the height of the tree which in turn decreases the length of all root-leaf paths by one.

So it maintains the **perfect balance** property

- i.e every path from the root to a leaf has the same length

This is the only way **Delete** can affect the length of paths so it also maintains the **perfect balance** property.

As each **Fuse** or **Transfer** takes $O(1)$ time, overall **Delete** takes $O(\log n)$ time.
The **DELETE** operation

To perform **DELETE**(k) on a leaf *(we’ll deal with other nodes later)*

**Step 1:** Search for the key *k* using **FIND**(k).
- use **FUSE** and **TRANSFER** to convert 2-nodes as we go down

**Step 2:** If the leaf is a 3-node,
- delete *(x, k)*, converting it into a 2-node

**Step 3:** If the leaf is a 4-node,
- delete *(x, k)*, converting it into a 3-node

As each **FUSE** or **TRANSFER** takes \(O(1)\) time, overall **DELETE** takes \(O(\log n)\) time
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?
What if we want to \textbf{DELETE} something other than a leaf?

\begin{itemize}
  \item \textbf{Step 1:} Find the \textbf{PREDECESSOR} of $k$ (this is essentially the same as \textbf{FIND})
  \begin{itemize}
    \item that’s the element with the largest key $k'$ such that $k' < k$
  \end{itemize}
\end{itemize}
What if we want to \textbf{DELETE} something other than a leaf?

\textbf{Step 1:} Find the \textbf{PREDECESSOR} of $k$ (this is essentially the same as \textbf{FIND})
- that’s the element with the largest key $k'$ such that $k' < k$
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$
What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$
What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of \( k \) (this is essentially the same as **FIND**)  
- that’s the element with the largest key \( k' \)  
such that \( k' < k \)
What if we want to delete something other than a leaf?

**Step 1:** Find the **predecessor** of $k$ (this is essentially the same as **find**)
- that’s the element with the largest key $k'$ such that $k' < k$

**Step 2:** Call **delete**($k'$)
- fortunately $k'$ is always a leaf
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$

**Step 2:** Call **DELETE**($k'$)
- fortunately $k'$ is *always* a leaf

**DELETE**($11$)

10 is the predecessor of 11
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$

**Step 2:** Call **DELETE**($k'$)
- fortunately $k'$ is *always* a leaf

**Step 3:** Overwrite $k$ with another copy of $k'$
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of \( k \) (this is essentially the same as **FIND**)  
- that’s the element with the largest key \( k' \) such that \( k' < k \)

**Step 2:** Call **DELETE**\((k')\)  
- fortunately \( k' \) is **always** a leaf

**Step 3:** Overwrite \( k \) with another copy of \( k' \)
What if we want to DELETE something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of \( k \) (this is essentially the same as **FIND**)
- that’s the element with the largest key \( k' \)
  such that \( k' < k \)

**Step 2:** Call **DELETE**\((k')\)
- fortunately \( k' \) is *always* a leaf

**Step 3:** Overwrite \( k \) with another copy of \( k' \)
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of \( k \) (this is essentially the same as **FIND**)
- that’s the element with the largest key \( k' \) such that \( k' < k \)

**Step 2:** Call **DELETE**\( (k') \)
- fortunately \( k' \) is *always* a leaf

**Step 3:** Overwrite \( k \) with another copy of \( k' \)
The **DELETE** operation

What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of $k$ (this is essentially the same as **FIND**)
- that’s the element with the largest key $k'$ such that $k' < k$

**Step 2:** Call **DELETE**($k'$)
- fortunately $k'$ is always a leaf

**Step 3:** Overwrite $k$ with another copy of $k'$
What if we want to **DELETE** something other than a leaf?

**Step 1:** Find the **PREDECESSOR** of \( k \) (this is essentially the same as **FIND**)
- that’s the element with the largest key \( k' \)
  such that \( k' < k \)

**Step 2:** Call **DELETE**\((k')\)
- fortunately \( k' \) is *always* a leaf

**Step 3:** Overwrite \( k \) with another copy of \( k' \)
  This also takes \( O(\log n) \) time
2-3-4 tree summary

A 2-3-4 is a data structure based on a tree structure which supports \textsc{Insert}(x, k), \textsc{Find}(k) and \textsc{Delete}(k)

each of these operations takes \textit{worst case} $O(\log n)$ time
2-3-4 tree summary

A 2-3-4 is a data structure based on a tree structure which supports \text{INSERT}(x, k), \text{FIND}(k) and \text{DELETE}(k)

each of these operations takes worst case \(O(\log n)\) time

Unfortunately, 2-3-4 trees are awkward to implement because the nodes don’t all have the same number of children
A 2-3-4 is a data structure based on a tree structure which supports \texttt{INSERT}(x, k), \texttt{FIND}(k) and \texttt{DELETE}(k) each of these operations takes \textit{worst case} $O(\log n)$ time.

Unfortunately, 2-3-4 trees are awkward to implement because the nodes don't all have the same number of children. So, what is used in practice?
Red-Black tree summary

A Red-Black tree is a data structure based on a binary tree structure which supports $\text{INSERT}(x, k)$, $\text{FIND}(k)$ and $\text{DELETE}(k)$ each of these operations takes worst case $O(\log n)$ time
Red-Black tree summary

A Red-Black tree is a data structure based on a binary tree structure which supports \( \text{INSERT}(x, k) \), \( \text{FIND}(k) \) and \( \text{DELETE}(k) \)

The root is black

each of these operations takes worst case \( O(\log n) \) time
Red-Black tree summary

A Red-Black tree is a data structure based on a **binary** tree structure which supports **INSERT** \( (x, k) \), **FIND** \( (k) \) and **DELETE** \( (k) \)

![Red-Black Tree Diagram]

Each of these operations takes **worst case** \( O(\log n) \) time

The root is **black**

All root-to-leaf paths have the same number of **black** nodes
Red-Black tree summary

A Red-Black tree is a data structure based on a binary tree structure which supports \textsc{Insert}(x, k), \textsc{Find}(k) and \textsc{Delete}(k)

Each of these operations takes worst case \(O(\log n)\) time

The root is \textbf{black}

All root-to-leaf paths have the same number of \textbf{black} nodes

\textbf{Red} nodes cannot have \textbf{red} children
Red-Black tree summary

A Red-Black tree is a data structure based on a binary tree structure which supports \textsc{insert}(x, k), \textsc{find}(k) and \textsc{delete}(k)

Each of these operations takes \textit{worst case} $O(\log n)$ time

The root is \textbf{black}

All root-to-leaf paths have the same number of \textbf{black} nodes

\textbf{Red} nodes cannot have \textbf{red} children

\textit{If these are used in practice, why did you waste our time with 2-3-4 trees?}
Red-Black tree summary

A Red-Black tree is a data structure based on a binary tree structure which supports \textsc{Insert}(x, k), \textsc{Find}(k) and \textsc{Delete}(k)

\begin{center}
\begin{tikzpicture}[level distance=1.5cm,sibling distance=1.5cm,thick]

  \node (root) {4}
  \node (left) {2} child {node (left-1) {1} \node (left-2) {3}}
  \node (right) {8} child {node (right-1) {6} child {node (right-1-1) {5}} child {node (right-1-2) {7}}}
  \node (right-2) {9}

\end{tikzpicture}
\end{center}

each of these operations takes \textit{worst case} $O(\log n)$ time

The root is \textbf{black}

All root-to-leaf paths have the same number of \textbf{black} nodes

\textbf{Red} nodes cannot have \textbf{red} children

\textit{If these are used in practice, why did you waste our time with 2-3-4 trees?}

1. 2-3-4 trees are conceptually much nicer
Red-Black tree summary

A Red-Black tree is a data structure based on a **binary** tree structure

which supports $\text{INSERT}(x, k)$, $\text{FIND}(k)$ and $\text{DELETE}(k)$

![Red-Black tree diagram]

each of these operations takes **worst case** $O(\log n)$ time

The root is **black**

All root-to-leaf paths have the same number of **black** nodes

**Red** nodes cannot have **red** children

*If these are used in practice, why did you waste our time with 2-3-4 trees?*

1. 2-3-4 trees are conceptually much nicer
2. they are secretly the same :)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree
(the details of the Red-Black tree operations are in CLRS Chapter 13)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree

(the details of the Red-Black tree operations are in CLRS Chapter 13)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree

(the details of the Red-Black tree operations are in CLRS Chapter 13)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree

(the details of the Red-Black tree operations are in CLRS Chapter 13)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree
(the details of the Red-Black tree operations are in CLRS Chapter 13)
2-3-4 trees vs. Red-Black trees

Any 2-3-4 tree can be converted into a Red-Black tree (and visa-versa)

The operations on 2-3-4 trees also have equivalent operations on a Red-Black tree

(the details of the Red-Black tree operations are in CLRS Chapter 13)
A **dynamic search structure** supports (at least) the following three operations:

- **DELETE**(\(k\)) - deletes the (unique) element \(x\) with \(x.key = k\)
- **INSERT**(\(x, k\)) - inserts \(x\) with key \(k = x.key\)
- **FIND**(\(k\)) - returns the (unique) element \(x\) with \(x.key = k\)

Here are the **worst case** time complexities of the structures we have seen...

<table>
<thead>
<tr>
<th>Structure</th>
<th>INSERT</th>
<th>DELETE</th>
<th>FIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsorted Linked List</td>
<td>(O(1))</td>
<td>(O(n))</td>
<td>(O(n))</td>
</tr>
<tr>
<td>Binary Search Tree</td>
<td>(O(n))</td>
<td>(O(n))</td>
<td>(O(n))</td>
</tr>
<tr>
<td>2-3-4 Tree</td>
<td>(O(\log n))</td>
<td>(O(\log n))</td>
<td>(O(\log n))</td>
</tr>
<tr>
<td>Red-Black Tree</td>
<td>(O(\log n))</td>
<td>(O(\log n))</td>
<td>(O(\log n))</td>
</tr>
</tbody>
</table>
End of part one
Part two
Skip lists

inspired by slides by Ashley Montanaro
Dynamic Search Structures

A **dynamic search structure**, stores a set of elements

*Each element* $x$ *must have a unique key* - $x.key$

The following operations are supported:

- **INSERT**($x, k$) - inserts $x$ with key $k = x.key$

- **FIND**($k$) - returns the (unique) element $x$ with $x.key = k$
  *(or reports that it doesn’t exist)*

- **DELETE**($k$) - deletes the (unique) element $x$ with $x.key = k$
  *(or reports that it doesn’t exist)*

We would also like it to support:

- **PREDECESSOR**($k$) - returns the (unique) element $x$
  with the largest key such that $x.key < k$

- **RANGEFIND**($k_1, k_2$) - returns every element $x$ with $k_1 \leq x.key \leq k_2$
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a sorted Linked List?

The bottleneck is **FIND**, which is *very inefficient*,
- we have to look through the entire linked list to find an item *(in the worst case)*

**INSERT** and **DELETE** also take $O(n)$ time *but only because they rely on FIND*

How can we speed up the **FIND** operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a *sorted* Linked List?

The bottleneck is \textsc{Find}, which is very inefficient,
- we have to look through the entire linked list to find an item \textit{(in the worst case)}

\textsc{Insert} and \textsc{Delete} also take $O(n)$ time \textit{but only because they rely on} \textsc{Find}

How can we speed up the \textsc{Find} operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a sorted Linked List?

The bottleneck is \texttt{FIND}, which is very inefficient,
- we have to look through the entire linked list to find an item \texttt{FIND}(18) \hspace{1cm} (in the worst case)

\texttt{INSERT} and \texttt{DELETE} also take \(O(n)\) time \textit{but only because they rely on \texttt{FIND}}

How can we speed up the \texttt{FIND} operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a *sorted* Linked List?

![Diagram of a sorted linked list with keys 1, 2, 5, 9, 16, 18, 25]

The bottleneck is **FIND**, which is very inefficient,
- we have to look through the entire linked list to find an item (in the worst case)

**INSERT** and **DELETE** also take $O(n)$ time *but only because they rely on FIND*

How can we speed up the **FIND** operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a sorted Linked List?

The bottleneck is FIND, which is very inefficient,
- we have to look through the entire linked list to find an item (in the worst case)

INSERT and DELETE also take $O(n)$ time but only because they rely on FIND

How can we speed up the FIND operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure

What about using a sorted Linked List?

The bottleneck is $\text{FIND}$, which is very inefficient,
- we have to look through the entire linked list to find an item (in the worst case)

$\text{INSERT}$ and $\text{DELETE}$ also take $O(n)$ time but only because they rely on $\text{FIND}$

How can we speed up the $\text{FIND}$ operation?
Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a *sorted* Linked List?

The bottleneck is **FIND**, which is *very inefficient*,
- we have to look through the entire linked list to find an item (in the worst case)

**INSERT** and **DELETE** also take \( O(n) \) time *but only because they rely on FIND*

How can we speed up the **FIND** operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure. What about using a *sorted* Linked List?

The bottleneck is **FIND**, which is *very inefficient*,
- we have to look through the entire linked list to find an item *(in the worst case)*

**INSERT** and **DELETE** also take $O(n)$ time *but only because they rely on FIND*

How can we speed up the **FIND** operation?
Using a Linked List as a Dynamic Search Structure (again)

Earlier we briefly considered using an unsorted Linked List as a dynamic search structure.

What about using a *sorted* Linked List?

The bottleneck is **FIND**, which is *very inefficient*,
- we have to look through the entire linked list to find an item *(in the worst case)*

**INSERT** and **DELETE** also take **$O(n)$** time *but only because they rely on FIND*

How can we speed up the **FIND** operation?
Making Shortcuts

How about adding some shortcuts?

1 → 2 → 5 → 9 → 16 → 18 → 25
Making Shortcuts

How about adding some shortcuts?

1 → 2 → 5 → 9 → 16 → 18 → 25
Making Shortcuts

How about adding some shortcuts?
Making Shortcuts

How about adding some shortcuts?
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform $\text{FIND}(k)$ we start in the top list
and go right until we come to a key $k' > k$

then we move down to the bottom list
and go right until we find $k$
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys.

To perform $\text{FIND}(k)$ we start in the top list and go right until we come to a key $k' > k$.

Then we move down to the bottom list and go right until we find $k$. 

\[
\begin{array}{c}
1 \rightarrow 2 \rightarrow 5 \rightarrow 9 \rightarrow 16 \rightarrow 18 \rightarrow 25 \\
1 \rightarrow 9 \rightarrow 25 \rightarrow \text{FIND}(18)
\end{array}
\]
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform \texttt{FIND}(k) we start in the top list and go right until we come to a key $k' > k$ then we move down to the bottom list and go right until we find $k$.
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform \texttt{FIND}(k) we start in the top list
and go right until we come to a key \( k' > k \)
then we move down to the bottom list
and go right until we find \( k \)
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform $\text{FIND}(k)$ we start in the top list
and go right until we come to a key $k' > k$
then we move down to the bottom list
and go right until we find $k$
How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform \textit{\textbf{FIND}}(k) we start in the top list

and go right until we come to a key \( k' > k \)

then we move down to the bottom list

and go right until we find \( k \)
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform \( \text{FIND}(k) \) we start in the top list
and go right until we come to a key \( k' > k \)
then we move down to the bottom list
and go right until we find \( k \)
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys…

To perform $\text{FIND}(k)$ we start in the top list
and go right until we come to a key $k' > k$
then we move down to the bottom list
and go right until we find $k$
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

To perform \( \text{FIND}(k) \) we start in the top list and go right until we come to a key \( k' > k \)

then we move down to the bottom list and go right until we find \( k \).
Making Shortcuts

How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys...

![Diagram]

To perform \( \text{FIND}(k) \) we start in the top list and go right until we come to a key \( k' > k \) then we move down to the bottom list and go right until we find \( k \).

How long does this take?
How about adding some shortcuts?

We’ve attached a second linked list containing only some of the keys…

To perform FIND(\(k\)) we start in the top list and go right until we come to a key \(k' > k\) and then we move down to the bottom list and go right until we find \(k\).

How long does this take? That depends on where we place the shortcuts.
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list...  
(the bottom list always contains all $n$ keys)
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list... 

*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a *Find* operation?
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list. . .

*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a *Find* operation?
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list...

\((the \ bottom \ list \ always \ contains \ all \ n \ keys)\)

where should we put them to minimise

the \textit{worst case} time for a \texttt{FIND} operation?

finding this is quick

finding this is slow
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list... 
*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a *Find* operation?
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list...

*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a **Find** operation?
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list...

*The bottom list always contains all \( n \) keys*

where should we put them to minimise

the *worst case* time for a FIND operation?

finding this is slow

finding this is quick
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list... (the bottom list always contains all $n$ keys)

where should we put them to minimise
the worst case time for a FIND operation?

![Diagram of linked lists with two levels]
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list...  
*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a FIND operation?
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list. . .

*(the bottom list always contains all \( n \) keys)*

where should we put them to minimise

the *worst case* time for a FIND operation?

If we spread out the \( m \) keys in the top list evenly . . .
Linked Lists with two levels

Imagine that we decide to place $m$ keys in the top list... 

*(the bottom list always contains all $n$ keys)*

where should we put them to minimise

the *worst case* time for a *FIND* operation?

If we spread out the $m$ keys in the top list evenly... 

the *worst case* time for a *FIND* operation becomes $O(m + n/m)$
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list . . .

\((\text{the bottom list always contains all } n \text{ keys})\)

where should we put them to minimise

the \textit{worst case} time for a \texttt{FIND} operation?

If we spread out the \( m \) keys in the top list evenly . . .

the \textit{worst case} time for a \texttt{FIND} operation becomes \( O(m + n/m) \)
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list... 

\((the \ bottom \ list \ always \ contains \ all \ n \ keys)\)

where should we put them to minimise 

the \textit{worst case} time for a FIND operation?

\[ \approx \frac{n}{m} \]

If we spread out the \( m \) keys in the top list evenly ... 

the \textit{worst case} time for a FIND operation becomes \( O(m + n/m) \)
Linked Lists with two levels

Imagine that we decide to place \( m \) keys in the top list . . .

(\textit{the bottom list always contains all} \( n \) \textit{keys})

where should we put them to minimise

the \textit{worst case} time for a \texttt{FIND} operation?

If we spread out the \( m \) keys in the top list evenly . . .

the \textit{worst case} time for a \texttt{FIND} operation becomes \( O(m + n/m) \)

By setting \( m = \sqrt{n} \), we get

the \textit{worst case} time for a \texttt{FIND} operation is \( O(\sqrt{n}) \)
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each **level** will now contain **half** of the keys *(rounding up)* from the level below

They are chosen to be as evenly spread as possible
Linked Lists with many levels

How about adding even more lists? (each list is called a level)

Each level will now contain \textbf{half} of the keys \textit{(rounding up)} from the level below.

They are chosen to be as evenly spread as possible.
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each **level** will now contain **half** of the keys *(rounding up)* from the level below

They are chosen to be as evenly spread as possible
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each *level* will now contain *half* of the keys *(rounding up)* from the level below.

They are chosen to be as evenly spread as possible.
Linked Lists with many levels

How about adding even more lists? (each list is called a level)

Each level will now contain half of the keys (rounding up) from the level below.
They are chosen to be as evenly spread as possible.
Linked Lists with many levels

How about adding even more lists? (each list is called a level)

Each level will now contain half of the keys (rounding up) from the level below

They are chosen to be as evenly spread as possible
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each **level** will now contain **half** of the keys *(rounding up)* from the level below

They are chosen to be as evenly spread as possible
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each *level* will now contain **half** of the keys *(rounding up)* from the level below

They are chosen to be as evenly spread as possible
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each *level* will now contain **half** of the keys *(rounding up)* from the level below.

They are chosen to be as evenly spread as possible.

The bottom level contains every key and every level contains the leftmost and rightmost keys.
Linked Lists with many levels

How about adding even more lists? *(each list is called a level)*

Each level will now contain **half** of the keys *(rounding up)* from the level below

They are chosen to be as evenly spread as possible

The bottom level contains every key

and every level contains the leftmost and rightmost keys

As each level contains half of the keys from the level below, there are $O(\log n)$ levels
**FIND** in multi-level linked lists

How do we perform **FIND**(\(k\)) in multi-level linked list?

*(essentially just like before)*
**FIND** in multi-level linked lists

How do we perform **FIND**\( (k) \) in multi-level linked list?

*(essentially just like before)*

To perform **FIND**\( (k) \),

1. Start at the top-left *(the head of the top level)*
2. While you haven’t found \( k \):
   - If the node to the right’s key, \( k' \leq k \)
     - Move right
   - Else
     - Move down
**FIND** in multi-level linked lists

How do we perform **FIND**($k$) in multi-level linked list?

*(essentially just like before)*

To perform **FIND**($k$),

- **Start at the top-left** (*the head of the top level*)
- While you haven’t found $k$:
  - If the node to the right’s key, $k' \leq k$
    - Move right
  - Else
    - Move down
How do we perform \texttt{FIND}(k) in multi-level linked list?

*(essentially just like before)*

To perform \texttt{FIND}(k),

Start at the top-left (the head of the top level)

While you haven’t found \( k \):

If the node to the right’s key, \( k' \leq k \)

Move right

Else

Move down
**FIND** in multi-level linked lists

How do we perform **FIND**($k$) in multi-level linked list?

*(essentially just like before)*

Start at the top-left (the head of the top level)

To perform **FIND**($k$),

While you haven’t found $k$:

If the node to the right’s key, $k′ \leq k$

Move right

Else

Move down

Consider **FIND**($35$)

38 > 35
**FIND** in multi-level linked lists

How do we perform $\text{FIND}(k)$ in multi-level linked list? (essentially just like before)

Start at the top-left (the head of the top level)

To perform $\text{FIND}(k)$,

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$
  Move right

Else
  Move down

Consider $\text{FIND}(35)$
To perform $\text{FIND}(k)$,

- **Start at the top-left (the head of the top level)**
- **While you haven’t found $k$:**
  - If the node to the right’s key, $k’ \leq k$
    - Move right
  - Else
    - Move down

How do we perform $\text{FIND}(k)$ in multi-level linked list?

(essentially just like before)
FIND in multi-level linked lists

How do we perform $\text{FIND}(k)$ in multi-level linked list?

(essentially just like before)

To perform $\text{FIND}(k)$,

Start at the top-left (the head of the top level)

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$

Move right

Else

Move down
How do we perform $\text{FIND}(k)$ in multi-level linked list?

(essentially just like before)

Start at the top-left (the head of the top level)

To perform $\text{FIND}(k)$,

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  - Move right

- Else
  - Move down
**FIND** in multi-level linked lists

How do we perform **FIND**(\(k\)) in multi-level linked list?

(essentially just like before)

Start at the top-left (the head of the top level)

To perform **FIND**(\(k\)),

While you haven’t found \(k\):

- If the node to the right’s key, \(k' \leq k\),
  - Move right
- Else
  - Move down
How do we perform $\text{FIND}(k)$ in multi-level linked list?

*essentially just like before*

To perform $\text{FIND}(k)$,

Start at the top-left (the head of the top level)

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  - Move right
- Else
  - Move down

Consider $\text{FIND}(35)$
**FIND** in multi-level linked lists

How do we perform **FIND**$(k)$ in multi-level linked list? 
*(essentially just like before)*

Start at the top-left *(the head of the top level)*

To perform **FIND**$(k)$,

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  - Move right
- Else
  - Move down

\[
\begin{align*}
38 > 35 & \quad \text{consider **FIND**}(35) \\
16 \leq 35 \\
25 \leq 35 \\
31 \leq 35 & \quad \text{consider **FIND**}(35) \\
\end{align*}
\]
FIND in multi-level linked lists

How do we perform FIND($k$) in multi-level linked list?

(essentially just like before)

Start at the top-left (the head of the top level)

To perform FIND($k$),

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  Move right

- Else
  Move down
How do we perform \texttt{FIND}(k) in multi-level linked list? 
(essentially just like before)

To perform \texttt{FIND}(k),

Start at the \textit{top-left} (the head of the top level)

While you haven’t found \( k \):

- If the node to the right’s key, \( k' \leq k \)
  - Move right
- Else
  - Move down
The complexity of \textsc{Find}

How long does \textsc{Find}(k) take in a multi-level linked list?
The complexity of *FIND*

How long does $\text{FIND}(k)$ take in a multi-level linked list?

**Observation 1** We only move down at most $O(\log n)$ times because there are only $O(\log n)$ levels.
The complexity of \textbf{FIND}

How long does \texttt{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only \texttt{move down} at most $O(\log n)$ times

\textit{because there are only $O(\log n)$ levels}

\textbf{Observation 2} Between any two nodes on level $i$,
there are at most 2 nodes on level $i + 1$

\textit{because we took half the nodes and spread them evenly}
The complexity of \textbf{FIND}

How long does \texttt{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most $O(\log n)$ times
because there are only $O(\log n)$ levels

\textbf{Observation 2} Between any two nodes on level $i$,
there are at most 2 nodes on level $i + 1$
because we took half the nodes and spread them evenly

\textbf{Observation 3} We only move right at most 2 times on any level $i + 1$
because we stopped moving right on level $i$
The complexity of \texttt{FIND}

How long does \texttt{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most $O(\log n)$ times.

because there are only $O(\log n)$ levels

\textbf{Observation 2} Between any two nodes on level $i$, there are at most 2 nodes on level $i + 1$.

because we took half the nodes and spread them evenly

\textbf{Observation 3} We only move right at most 2 times on any level $i + 1$.

because we stopped moving right on level $i$
The complexity of \textit{FIND}

How long does \textit{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most \(O(\log n)\) times

\textit{because there are only} \(O(\log n)\) \textit{levels}

\textbf{Observation 2} Between any two nodes on level \(i\),

\textit{there are at most 2 nodes on level} \(i + 1\)

\textit{because we took half the nodes and spread them evenly}

\textbf{Observation 3} We only move right at most 2 times on any level \(i + 1\)

\textit{because we stopped moving right on level} \(i\)
The complexity of \textbf{FIND}

How long does \textbf{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most $O(\log n)$ times

\textit{because there are only} $O(\log n)$ \textit{levels}

\textbf{Observation 2} Between any two nodes on level $i$,
\textit{there are at most 2 nodes on level $i+1$}

\textit{because we took half the nodes and spread them evenly}

\textbf{Observation 3} We only move right at most 2 times on any level $i+1$

\textit{because we stopped moving right on level $i$}
The complexity of \texttt{FIND}

How long does \texttt{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most \(O(\log n)\) times 

\textit{because there are only \(O(\log n)\) levels}

\textbf{Observation 2} Between any two nodes on level \(i\), 
   there are at most 2 nodes on level \(i + 1\) 

\textit{because we took half the nodes and spread them evenly}

\textbf{Observation 3} We only move right at most 2 times on any level \(i + 1\) 

\textit{because we stopped moving right on level \(i\)}
The complexity of \texttt{FIND}

How long does \texttt{FIND}(k) take in a multi-level linked list?

\textbf{Observation 1} We only move down at most \(O(\log n)\) times

\textit{because there are only} \(O(\log n)\) \textit{levels}

\textbf{Observation 2} Between any two nodes on level \(i\),
there are at most 2 nodes on level \(i + 1\)

\textit{because we took half the nodes and spread them evenly}

\textbf{Observation 3} We only move right at most 2 times on any level \(i + 1\)

\textit{because we stopped moving right on level} \(i\)
The complexity of **FIND**

How long does **FIND**(*k*) take in a multi-level linked list?

**Observation 1** We only move down at most **O**(log *n*) times.

because there are only **O**(log *n*) levels

**Observation 2** Between any two nodes on level *i*,
there are at most 2 nodes on level *i* + 1.

because we took half the nodes and spread them evenly

**Observation 3** We only move right at most 2 times on any level *i* + 1.

because we stopped moving right on level *i*
The complexity of $\text{FIND}$

How long does $\text{FIND}(k)$ take in a multi-level linked list?

**Observation 1** We only **move down** at most $O(\log n)$ times  
*because there are only $O(\log n)$ levels*

**Observation 2** Between any two nodes on level $i$, there are **at most 2 nodes** on level $i + 1$ *because we took half the nodes and spread them evenly*

**Observation 3** We only **move right** at most 2 times on any level $i + 1$ *because we stopped moving right on level $i$*

**Fact** We only **move** at most $O(\log n)$ times while performing a $\text{FIND}$
Multi-level Linked Lists

If we had a multi-level linked list with $O(\log n)$ levels

where each level contained half of the keys from the level below

and the keys were evenly spread as possible

then we could perform FIND in $O(\log n)$ time
Multi-level Linked Lists

If we had a multi-level linked list with $O(\log n)$ levels

where each level contained half of the keys from the level below

and the keys were evenly spread as possible

then we could perform FIND in $O(\log n)$ time

How are we going to do INSERTS and DELETES?
If we had a multi-level linked list with $O(\log n)$ levels where each level contained half of the keys from the level below and the keys were evenly spread as possible, then we could perform FIND in $O(\log n)$ time.

How are we going to do INSERTS and DELETES? Which levels should we put an INSERTED key into?
Multi-level Linked Lists

If we had a multi-level linked list with \( O(\log n) \) levels

where each level contained half of the keys from the level below

and the keys were evenly spread as possible

then we could perform FIND in \( O(\log n) \) time

How are we going to do INSERTS and DELETES?

Which levels should we put an INSERTED key into?

How can we keep a good spread of keys at each levels?
Multi-level Linked Lists

If we had a multi-level linked list with $O(\log n)$ levels, where each level contained half of the keys from the level below and the keys were evenly spread as possible, then we could perform FIND in $O(\log n)$ time.

How are we going to do INSERTS and DELETES? Which levels should we put an INSERTED key into? How can we keep a good spread of keys at each level? especially when we don’t know what will be INSERTED and DELETED in the future.
Multi-level Linked Lists

If we had a multi-level linked list with $O(\log n)$ levels
where each level contained half of the keys from the level below
and the keys were evenly spread as possible

then we could perform FIND in $O(\log n)$ time

If you can’t get organised, get randomised

How are we going to do INSERTS and DELETES?
Which levels should we put an INSERTED key into?
How can we keep a good spread of keys at each level?
especially when we don’t know what will be INSERTED and DELETED in the future
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level

Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level

Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level
Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level
Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level

Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...
For each key that got a head, put it in the new top level
Repeat with the keys from the new top level
(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level
Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let's rewind and try building another Multi-level Linked List... by flipping coins.

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level
Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...

For each key that got a head, put it in the new top level

Repeat with the keys from the new top level

(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins.

(we still always include the smallest and largest keys in every level)

Flip one coin for each key...
For each key that got a head, put it in the new top level
Repeat with the keys from the new top level
(stop when the top level contains only the smallest and largest keys)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)
Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins.

(we still always include the smallest and largest keys in every level)

This doesn’t look quite perfect but actually, it’s very good with high probability (more on this later)
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

This doesn’t look quite perfect but actually, it’s very good with high probability (more on this later)

The intuition is that \( n \) coin flips contain about \( \frac{n}{2} \) heads and about \( \frac{n}{2} \) tails
Building Multi-level Linked Lists by flipping coins

Before we formally introduce Skip Lists, we let’s rewind and try building another Multi-level Linked List... by flipping coins

(we still always include the smallest and largest keys in every level)

This doesn’t look quite perfect but actually, it’s very good with high probability (more on this later)

The intuition is that $n$ coin flips contain about $\frac{n}{2}$ heads and about $\frac{n}{2}$ tails and the heads are roughly evenly spread out
Skip Lists

A skip list is a multi-level linked list where

the **INSERTS** are done by flipping coins

*i.e. this is a skip list.*
Skip Lists

A skip list is a multi-level linked list where
the **INSERTS** are done by flipping coins
i.e. this is a skip list...

To perform \text{INSERT}(x, k),

**Step 1:** Use \text{FIND}(k) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert \((x, k)\) into the next level up

* (if there is no ‘next level up’, create a new level at the top)

If you get a **tails**, **stop**
Skip Lists

A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list...*

To perform **INSERT**(\(x, k\)),

**Step 1:** Use **FIND**(\(k\)) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert \((x, k)\) into the next level up

*if there is no ‘next level up’, create a new level at the top*

If you get a **tails**, **stop**
Skip Lists

A skip list is a multi-level linked list where
the **INSERTS** are done by flipping coins

*i.e. this is a skip list...*

To perform **INSERT**($x, k$),

**Step 1:** Use **FIND**($k$) to insert $(x, k)$ into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert $(x, k)$ into the next level up
  
  *if there is no ‘next level up’, create a new level at the top*

- If you get a **tails**, **stop**
Skip Lists

A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

\[ \text{i.e. this is a skip list...} \]

To perform **INSERT**(\(x, k\)),

**Step 1**: Use **FIND**(\(k\)) to insert \((x, k)\) into the bottom level

**Step 2**: Flip a coin repeatedly:

- **If** you get a **heads**, insert \((x, k)\) into the next level up
  
  *(if there is no ‘next level up’, create a new level at the top)*

- **If** you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list* . . .

To perform **INSERT** \((x, k)\),

**Step 1**: Use **FIND** \(k\) to insert \((x, k)\) into the bottom level

**Step 2**: Flip a coin repeatedly:

If you get a heads, insert \((x, k)\) into the next level up

*(if there is no ‘next level up’, create a new level at the top)*

If you get a tails, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

* i.e. this is a skip list...

To perform **INSERT**(*x, k*),

**Step 1:** Use **FIND**(*k*) to insert (**x, k**) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert (**x, k**) into the next level up

* (if there is no ‘next level up’, create a new level at the top)*

If you get a **tails**, **stop**
A skip list is a multi-level linked list where the \textbf{INSERTS} are done by flipping coins \emph{i.e. this is a skip list}.

To perform $\text{INSERT}(x, k)$,

**Step 1:** Use $\text{FIND}(k)$ to insert $(x, k)$ into the bottom level

**Step 2:** Flip a coin repeatedly:
- If you get a \textbf{heads}, insert $(x, k)$ into the next level up \\
  (if there is no ‘next level up’, create a new level at the top)
- If you get a \textbf{tails}, stop
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins. *i.e. this is a skip list.*

To perform **INSERT**\((x, k)\),

**Step 1:** Use **FIND**\((k)\) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert \((x, k)\) into the next level up
  
  *(if there is no ‘next level up’, create a new level at the top)*

- If you get a **tails**, **stop**
Skip Lists

A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

* i.e. this is a skip list...

To perform **INSERT**\((x, k)\),

**Step 1:** Use **FIND**\((k)\) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert \((x, k)\) into the next level up

*(if there is no 'next level up', create a new level at the top)*

If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list.*

To perform **\( \text{INSERT}(x, k) \)**,

**Step 1:** Use \( \text{FIND}(k) \) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert \((x, k)\) into the next level up
- If you get a **tails**, stop

*(if there is no ‘next level up’, create a new level at the top)*

19 goes in here
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins, *i.e. this is a skip list*.

To perform **INSERT**(*x*, *k*),

**Step 1:** Use **FIND**(*k*) to insert (*x*, *k*) into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert (*x*, *k*) into the next level up

  *(if there is no ‘next level up’, create a new level at the top)*

- If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list...*

To perform **INSERT**(*x, k*),

**Step 1:** Use **FIND**(*k*) to insert (*x, k*) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert (*x, k*) into the next level up

*if there is no ‘next level up’, create a new level at the top*

If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list.*

To perform **INSERT**($x, k$),

**Step 1:** Use **FIND**($k$) to insert $(x, k)$ into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert $(x, k)$ into the next level up
  
  *(if there is no ‘next level up’, create a new level at the top)*

- If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERT**s are done by flipping coins *i.e. this is a skip list.*

To perform **INSERT**\((x, k)\),

**Step 1:** Use **FIND**\((k)\) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert \((x, k)\) into the next level up *(if there is no ‘next level up’, create a new level at the top)*
- If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins

*i.e. this is a skip list.*

To perform **INSERT**(\(x, k\)),

**Step 1:** Use **FIND**(\(k\)) to insert \((x, k)\) into the bottom level

**Step 2:** Flip a coin repeatedly:

If you get a **heads**, insert \((x, k)\) into the next level up

*(if there is no ‘next level up’, create a new level at the top)*

If you get a **tails**, stop
Skip Lists

A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins, *i.e. this is a skip list*...

To perform **INSERT**($x, k$),

**Step 1:** Use **FIND**($k$) to insert $(x, k)$ into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert $(x, k)$ into the next level up
  *(if there is no ‘next level up’, create a new level at the top)*

- If you get a **tails**, **stop**
A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins *i.e. this is a skip list...*

To perform $$\text{INSERT}(x, k),$$

**Step 1:** Use $$\text{FIND}(k)$$ to insert $$(x, k)$$ into the bottom level

**Step 2:** Flip a coin repeatedly:

- If you get a **heads**, insert $$(x, k)$$ into the next level up  
  *(if there is no 'next level up', create a new level at the top)*
- If you get a **tails**, **stop**
Skip Lists

A skip list is a multi-level linked list where the **INSERTS** are done by flipping coins, i.e. this is a skip list...

To perform **INSERT**(\(x, k\)),

**Step 1**: Use **FIND**(\(k\)) to insert \((x, k)\) into the bottom level

**Step 2**: Flip a coin repeatedly:

- If you get a **heads**, insert \((x, k)\) into the next level up
  (if there is no ‘next level up’, create a new level at the top)
- If you get a **tails**, **stop**
That about **DELETES**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels.
That about **DELETES**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels.

To perform **DELETE**($k$),

**Step 1:** Use **FIND**($k$) to find ($x$, $k$)

**Step 2:** Delete ($x$, $k$) from all levels
That about **DELETE**s?

**DELETEING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**($k$),

**Step 1:** Use **FIND**($k$) to find ($x$, $k$)

**Step 2:** Delete ($x$, $k$) from all levels
That about **DELETES**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**\((k)\),

**Step 1:** Use **FIND**\((k)\) to find \((x, k)\)

**Step 2:** Delete \((x, k)\) from all levels
That about **DELETE**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**($k$),

**Step 1:** Use **FIND**($k$) to find ($x$, $k$)

**Step 2:** Delete ($x$, $k$) from all levels
That about **DELETES**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**($k$),

**Step 1**: Use **FIND($k$)** to find ($x, k$)

**Step 2**: Delete ($x, k$) from all levels
That about **DELETE**s?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**(\(k\)),

**Step 1:** Use **FIND**(\(k\)) to find \((x, k)\)

**Step 2:** Delete \((x, k)\) from all levels
That about **DELETE**s?

**Deleting** is straightforward, just **FIND** the key and **DELETE** it from all levels.

To perform **DELETE**\((k)\),

**Step 1:** Use **FIND**\((k)\) to find \((x, k)\)

**Step 2:** Delete \((x, k)\) from all levels.
That about **DELETES**?

**DELETING** is straightforward, just find the key and **DELETE** it from all levels.

To perform **DELETE**($k$),

**Step 1**: Use **FIND**($k$) to find ($x$, $k$)

**Step 2**: Delete ($x$, $k$) from all levels
That about **DELETE**s?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels.

To perform **DELETE**(*k*),

**Step 1**: Use **FIND**(*k*) to find (*x*, *k*)

**Step 2**: Delete (*x*, *k*) from all levels

**Step 3**: Remove any empty levels

*(ones containing only the smallest and largest keys)*
That about **DELETE**?

**DELETING** is straightforward, just **FIND** the key and **DELETE** it from all levels

To perform **DELETE**(\(k\)),

**Step 1:** Use **FIND**(\(k\)) to find \((x, k)\)

**Step 2:** Delete \((x, k)\) from all levels

**Step 3:** Remove any empty levels

(ones containing only the smallest and largest keys)
Skip Lists

That about DELETE?

DELETE is straightforward, just FIND the key and DELETE it from all levels

To perform DELETE($k$),

Step 1: Use FIND($k$) to find ($x$, $k$)

Step 2: Delete ($x$, $k$) from all levels

Step 3: Remove any empty levels

(ones containing only the smallest and largest keys)
A skip list is a randomised data structure, based on link lists with shortcuts which supports \textsc{Insert}(x, k), \textsc{Find}(k) and \textsc{Delete}(k)

We will show that each of these operations takes expected $O(\log n)$ time

That is, they take $O(\log n)$ time ‘on average’

\textbf{Important} There is \textit{no randomness in the data},

\textit{the only randomness is in the coin flips}

On the worst case input sequence, the expected time is $O(\log n)$
How many levels are in a Skip list?

We begin by proving that after \( n \) \textsc{insert} operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...
How many levels are in a Skip list?

We begin by proving that after $n$ \textsc{insert} operations, a skip list is very unlikely to have more than $2 \log n$ levels...

An empty skip list contains only one level
and the only way this can increase is during an \textsc{insert} operation
How many levels are in a Skip list?

We begin by proving that after \( n \) INSERT operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some INSERT(\( x, k \)) operation.
How many levels are in a Skip list?

We begin by proving that after $n$ INSERT operations, a skip list is very unlikely to have more than $2 \log n$ levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some $\text{INSERT}(x, k)$ operation.

The probability $(x, k)$ is inserted into more than 1 level is $\frac{1}{2}$

*(the first coin flip is H)*
How many levels are in a Skip list?

We begin by proving that after \( n \) \textsc{Insert} operations, a skip list is very unlikely to have more than \( 2 \log n \) levels.

An empty skip list contains only one level and the only way this can increase is during an \textsc{Insert} operation.

Consider some \textsc{Insert}(x, k) operation

The probability \((x, k)\) is inserted into more than 1 level is \(\frac{1}{2}\) 
\((\text{the first coin flip is H})\)

The probability \((x, k)\) is inserted into more than 2 levels is \(\frac{1}{4}\) 
\((\text{we throw HH...})\)
How many levels are in a Skip list?

We begin by proving that after \( n \) INSERT operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some \text{INSERT}(x, k)\ operation

The probability \( (x, k) \) is inserted into more than 1 level is \( \frac{1}{2} \)

\( \text{(the first coin flip is H)} \)

The probability \( (x, k) \) is inserted into more than 2 levels is \( \frac{1}{4} \)

\( \text{(we throw HH...)} \)

The probability \( (x, k) \) is inserted into more than 3 levels is \( \frac{1}{8} \)

\( \text{(we throw HHH...)} \)
How many levels are in a Skip list?

We begin by proving that after $n$ `INSERT` operations, a skip list is very unlikely to have more than $2 \log n$ levels...

An empty skip list contains only one level and the only way this can increase is during an `INSERT` operation

Consider some `INSERT(x, k)` operation

The probability $(x, k)$ is inserted into more than 1 level is $\frac{1}{2}$

*(the first coin flip is $H$)*

The probability $(x, k)$ is inserted into more than 2 levels is $\frac{1}{4}$

*(we throw $HH$...)*

The probability $(x, k)$ is inserted into more than 3 levels is $\frac{1}{8}$

*(we throw $HHH$...)*

The probability $(x, k)$ is inserted into more than $j$ levels is $\frac{1}{2^j}$
How many levels are in a Skip list?

We begin by proving that after $n$ INSERT operations, a skip list is very unlikely to have more than $2 \log n$ levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some INSERT$(x, k)$ operation.

The probability $(x, k)$ is inserted into more than $j$ levels is $\frac{1}{2^j}$. 
How many levels are in a Skip list?

We begin by proving that after $n$ INSERT operations, a skip list is very unlikely to have more than $2 \log n$ levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some $\text{INSERT}(x, k)$ operation.

The probability $(x, k)$ is inserted into more than $2 \log n$ levels is $\frac{1}{2^{2 \log n}}$. 
How many levels are in a Skip list?

We begin by proving that after \( n \) INSERT operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some INSERT\((x, k)\) operation

The probability \((x, k)\) is inserted into more than \(2 \log n\) levels is \(\frac{1}{2^{2 \log n}} = \frac{1}{n^2}\)
How many levels are in a Skip list?

We begin by proving that after \( n \) \textsc{insert} operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...

An empty skip list contains only one level and the only way this can increase is during an \textsc{insert} operation.

Consider some \textsc{insert}(\( x, k \)) operation

The probability \((x, k)\) is inserted into more than \( 2 \log n \) levels is

\[
\frac{1}{2^{2 \log n}} = \frac{1}{n^2}
\]

The \textit{union} bound

Let \( E_1, E_2 \ldots E_n \) be events where \( E_j \) occurs with probability \( p_j \).

The probability of at least one \( E_j \) occurring is at most \( \sum_j p_j \).
How many levels are in a Skip list?

We begin by proving that after $n$ \texttt{INSERT} operations, a skip list is very unlikely to have more than $2 \log n$ levels.

An empty skip list contains only one level and the only way this can increase is during an \texttt{INSERT} operation.

Consider some \texttt{INSERT}(\texttt{x, k}) operation.

The probability (\texttt{x, k}) is inserted into more than $2 \log n$ levels is $\frac{1}{2^{2 \log n}} = \frac{1}{n^2}$

**The union bound**

Let $E_1, E_2 \ldots E_n$ be events where $E_j$ occurs with probability $p_j$.

The probability of at least one $E_j$ occurring is at most $\sum_j p_j$.

Let $E_j$ be the event that the $j$-th \texttt{INSERT} puts its element in more than $2 \log n$ levels.
How many levels are in a Skip list?

We begin by proving that after \( n \) `INSERT` operations, a skip list is very unlikely to have more than \( 2 \log n \) levels...

An empty skip list contains only one level
and the only way this can increase is during an `INSERT` operation

Consider some `INSERT(x, k)` operation

The probability \((x, k)\) is inserted into more than \(2 \log n\) levels is

\[
\frac{1}{2^{2 \log n}} = \frac{1}{n^2}
\]

The **union** bound

Let \(E_1, E_2 \ldots E_n\) be events where \(E_j\) occurs with probability \(p_j\)

The probability of at least one \(E_j\) occurring is at most \(\sum_j p_j\)

Let \(E_j\) be the event that the \(j\)-th `INSERT` puts its element in more than \(2 \log n\) levels

The probability of at least one \(E_j\) occurring is at most \(\sum_j \frac{1}{n^2}\)
How many levels are in a Skip list?

We begin by proving that after \( n \) INSERT operations, a skip list is very unlikely to have more than \( 2 \log n \) levels.

An empty skip list contains only one level and the only way this can increase is during an INSERT operation.

Consider some \( \text{INSERT}(x, k) \) operation.

The probability \( (x, k) \) is inserted into more than \( 2 \log n \) levels is

\[
\frac{1}{2^{2 \log n}} = \frac{1}{n^2}
\]

**The union bound**

Let \( E_1, E_2 \ldots E_n \) be events where \( E_j \) occurs with probability \( p_j \).

The probability of at least one \( E_j \) occurring is at most

\[
\sum_j p_j
\]

Let \( E_j \) be the event that the \( j \)-th INSERT puts its element in more than \( 2 \log n \) levels.

The probability of at least one \( E_j \) occurring is at most

\[
\sum_j \frac{1}{n^2} = \frac{1}{n}
\]
How many levels are in a Skip list?

After \( n \) INSERT operations, the probability that a skip list has more than \( 2 \log n \) levels... is at most \( \frac{1}{n} \).
How many levels are in a Skip list?

After \( n \) INSERT operations, the probability that a skip list has more than \( 2 \log n \) levels... is at most \( \frac{1}{n} \)

*It gets better as \( n \) increases!*
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),

we can conclude that the number of times we **move down** is very likely to be $O(\log n)$

Start at the top-left (the head of the top level)

To perform **FIND**(k),

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  - Move right

- Else
  - Move down
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),

we can conclude that the number of times we **move down** is very likely to be $O(\log n)$

but how many times do we **move right**?

To perform **FIND**($k$),

Start at the top-left (the head of the top level)

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$
    Move right

Else Move down
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$
but how many times do we move right?

Start at the top-left (the head of the top level)

To perform FIND($k$),
While you haven’t found $k$:
    If the node to the right’s key, $k' \leq k$
        Move right
    Else
        Move down
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we \textit{move down} is very likely to be $O(\log n)$

but how many times do we \textit{move right}? 

Start at the top-left (the head of the top level)

To perform $\texttt{FIND}(k)$,

\begin{itemize}
  \item While you haven’t found $k$:
    \begin{itemize}
      \item If the node to the right’s key, $k' \leq k$
        \begin{itemize}
          \item Move right
        \end{itemize}
      \item Else
        \begin{itemize}
          \item Move down
        \end{itemize}
    \end{itemize}
\end{itemize}

Consider $\texttt{FIND}(35)$
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Start at the top-left (the head of the top level)

To perform FIND($k$),

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$
Move right

Else Move down
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we **move down** is very likely to be $O(\log n)$

but how many times do we **move right**?

Start at the top-left (the head of the top level)

To perform **FIND**($k$),

While you haven’t found $k$:

1. If the node to the right’s key, $k' \leq k$ **Move right**
2. Else **Move down**

consider **FIND**($35$)
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),

we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Start at the top-left (the head of the top level)

To perform $\text{FIND}(k)$,

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$

Move right

Else Move down
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability), we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

To perform $\texttt{FIND}(k)$,

Start at the top-left \textit{(the head of the top level)}

While you haven’t found $k$:

If the node to the right’s key, $k' \leq k$
    Move right

Else Move down
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Start at the top-left (the head of the top level)

To perform FIND($k$),

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$
  - Move right

- Else
  - Move down
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability), we can conclude that the number of times we move down is very likely to be $O(\log n)$ but how many times do we move right?

Start at the top-left (the head of the top level)

To perform FIND($k$),

While you haven’t found $k$:

- If the node to the right’s key, $k’ \leq k$ Move right
- Else Move down

Consider FIND(35)
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability), we can conclude that the number of times we move down is very likely to be $O(\log n)$ but how many times do we move right?

Start at the top-left (the head of the top level)

To perform \texttt{FIND}(k),

\begin{itemize}
  \item If the node to the right’s key, $k' \leq k$:
    \begin{itemize}
      \item Move right
    \end{itemize}
  \item Else Move down
\end{itemize}
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability), we can conclude that the number of times we move down is very likely to be $O(\log n)$ but how many times do we move right?

---

Start at the top-left (the head of the top level)

To perform $\text{FIND}(k)$,

While you haven’t found $k$:

- If the node to the right’s key, $k' \leq k$, Move right
- Else Move down

consider $\text{FIND}(35)$
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is \( \mathcal{O}(\log n) \) (with high probability),
we can conclude that the number of times we \texttt{move down} is very likely to be \( \mathcal{O}(\log n) \)
but how many times do we \texttt{move right}?

Start at the top-left (the head of the top level)

To perform \texttt{FIND}(k),

While you haven’t found \( k \):

\begin{itemize}
  \item If the node to the right’s key, \( k' \leq k \)
    \begin{itemize}
      \item Move right
    \end{itemize}
  \item Else \ Move down
\end{itemize}
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),

we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Consider \texttt{FIND(35)}
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we **move down** is very likely to be $O(\log n)$
but how many times do we **move right**?

Consider **FIND**(35)

How long is this path?
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),

we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Consider **FIND**(35)

How long is this path?

1. Reverse it
So how long does a `FIND` take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$

but how many times do we move right?

Consider `FIND(35)`

How long is this path?

1. Reverse it
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$
but how many times do we move right?

Consider \texttt{FIND}(35)

How long is this path?

1. Reverse it
2. Convince yourself this is the same path:

Start at $k$
While not at the top-left:
   If you can, Move up
   Else Move left
So how long does a FIND take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we move down is very likely to be $O(\log n)$
but how many times do we move right?

How long is this path?
1. Reverse it
2. Convince yourself this is the same path:
3. Now convince yourself
   it takes the same time as this: (in expectation)

Start at $k$
While not at the top-left:
   If (flip a coin)
      Move up
   Else
      Move left
So how long does a **FIND** take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability), we can conclude that the number of times we move down is very likely to be $O(\log n)$ but how many times do we move right?

How long is this path?

1. Reverse it
2. Convince yourself this is the same path:
3. Now convince yourself it takes the same time as this: *(in expectation)*

Start at $k$
While not at the top-left:
   - If (flip a coin) Move up
   - Else Move left
So how long does a \texttt{FIND} take? (sketch proof)

As the number of levels is $O(\log n)$ (with high probability),
we can conclude that the number of times we \texttt{move down} is very likely to be $O(\log n)$
but how many times do we \texttt{move right}? $O(\log n)$ in expectation

How long is this path?

1. Reverse it
2. Convince yourself this is the same path:
3. Now convince yourself
   it takes the same time as this:
   \textit{(in expectation)}

Start at $k$
While not at the top-left:
\begin{align*}
\text{If (flip a coin) } & \text{Move up} \\
\text{Else } & \text{Move left}
\end{align*}
Time complexities

When performing a \texttt{FIND} operation, the number of moves is $O(\log n)$ in expectation.
Time complexities

When performing a **FIND** operation, the number of moves is $O(\log n)$ in expectation as each move takes $O(1)$ time, the expected time complexity is $O(\log n)$
Time complexities

When performing a \texttt{FIND} operation, the number of moves is $O(\log n)$ in expectation as each move takes $O(1)$ time, the expected time complexity is $O(\log n)$.

The number of levels is also $O(\log n)$ in expectation.
Time complexities

When performing a FIND operation, the number of moves is $O(\log n)$ in expectation as each move takes $O(1)$ time, the expected time complexity is $O(\log n)$.

The number of levels is also $O(\log n)$ in expectation.

Both INSERT and DELETE also take expected $O(\log n)$ time. This is because they both call FIND and then spend $O(1)$ time per level.
Time complexities

When performing a **FIND** operation, the number of moves is $O(\log n)$ in expectation as each move takes $O(1)$ time, the expected time complexity is $O(\log n)$

The number of levels is also $O(\log n)$ in expectation

Both **INSERT** and **DELETE** also take expected $O(\log n)$ time

*this is because they both call **FIND** and then spend $O(1)$ time per level*

In fact, all three operations actually take $O(\log n)$ time

*with high probability*

i.e. the probability of an operation taking longer is at most $\frac{1}{n}$
Time complexities

When performing a **FIND** operation, the number of moves is $O(\log n)$ in expectation as each move takes $O(1)$ time, the expected time complexity is $O(\log n)$

The number of levels is also $O(\log n)$ in expectation

Both **INSERT** and **DELETE** also take expected $O(\log n)$ time

  *this is because they both call **FIND** and then spend $O(1)$ time per level*

In fact, all three operations actually take $O(\log n)$ time

  *with high probability*

  i.e. the probability of an operation taking longer is at most $\frac{1}{n}$

  *(this is a stronger claim but proving it is harder)*
Skip Lists (post-proof) summary

A skip list is a randomised data structure, based on link lists with shortcuts which supports \texttt{INSERT}(x, k), \texttt{FIND}(k) and \texttt{DELETE}(k)

Each of these operations takes expected $O(\log n)$ time.

That is, they take $O(\log n)$ time ‘on average’.

\textbf{Important} There is \textit{no randomness in the data},
the only randomness is in the coin flips.

On the worst case input sequence, the expected time is $O(\log n)$.
Dynamic Search Structure Summary

A dynamic search structure supports (at least) the following three operations

\[
\begin{align*}
\text{DELETE}(k) & \quad - \text{ deletes the (unique) element } x \text{ with } x.\text{key} = k \\
\text{INSERT}(x, k) & \quad - \text{ inserts } x \text{ with key } k = x.\text{key} \\
\text{FIND}(k) & \quad - \text{ returns the (unique) element } x \text{ with } x.\text{key} = k
\end{align*}
\]

Here are the time complexities of the structures we have seen...

<table>
<thead>
<tr>
<th></th>
<th>INSERT</th>
<th>DELETE</th>
<th>FIND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsorted Linked List</td>
<td>$O(1)$</td>
<td>$O(n)$</td>
<td>$O(n)$</td>
</tr>
<tr>
<td>Binary Search Tree</td>
<td>$O(n)$</td>
<td>$O(n)$</td>
<td>$O(n)$</td>
</tr>
<tr>
<td>2-3-4 Tree</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
</tr>
<tr>
<td>Red-Black Tree</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
</tr>
<tr>
<td>Skip list</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
<td>$O(\log n)$</td>
</tr>
</tbody>
</table>

The time complexities for the Skip list are *expected*, for the others, they are *worst case*