CIS 451/551: B+ tree insertion

related to a homework 5 problem
text’s rules (corrected hopefully)

• **internal node** of order $p$ structure:
  $$<P_1, K_1, P_2, K_2, ..., K_{q-1}, P_q>$$

• $\text{ceil}(p/2) \leq q \leq p$

• **leaf node** of order $p_{\text{leaf}}$:
  $$< (K_1, P_1), (K_2, P_2), ..., (K_q, P_q), P_{\text{next}} >$$

• $\text{ceil}(p_{\text{leaf}}/2) \leq q \leq p_{\text{leaf}}$  
  (←corrected part ??)

• usually $p_{\text{leaf}} = p - 1$
left bias on splitting

overflowing leaf node split ($p_{leaf}=4$)

(a b c d e)

becomes

(c)

(a b c)

(d e)

promote c by inserting into parent

overflowing internal node split ($p=4$)

(a b c d)

becomes

(c)

(a b)

(d)

similarly c gets promoted

promoted value remains in leaf node, does not remain in internal node

no “bias” if p or $p_{leaf}$ are odd
now the problem

insert into empty B+ tree with parameters p=4 and $p_{\text{leaf}}=3$ the values:

20, 40, 37, 27, 25, 30, 15, 60, 42, 35, 5, 22, 50, 11, 32
recall with $p=4$ and $p_{\text{leaf}}=3$

- leaf node can have 2-3 keys, with 2-3 child pointers
- internal node can have 1-3 keys with 2-4 pointers
- the root is an exception: it can have a single key
- we are using the B+ tree “rules” from our text
- other texts will have slightly different rules so beware
starting to insert

after 20: (20)

after 37: (20, 37, 40)

after 40: (20, 40)
the first split

after 27 overflow, so need to split since $p_{\text{leaf}}=3$:

(20, 27, 37, 40)

splits to

(27)

(20, 27)  (37, 40)

then inserting 25 and 30 is boring:

(27)

(20, 25, 25)  (30, 37, 40)
after 15 overflow again:

```
(27)
  /  \
(15, 20, 25, 25) (30, 37, 40)
```

splits

```
(20, 27)
  /  \
(15, 20) (25, 27) (30, 37, 40)
```

another after 60:

```
(20, 27)
  /  \
(15, 20) (25, 27) (30, 37, 40, 60)
```

becomes

```
(20, 27, 37)
  /  \
(15, 20) (25, 27) (30, 37) (40, 60)
```
now after 42, 35, 5, 22:

(5, 15, 20) (22, 25, 27) (30, 35, 37) (40, 42, 60)
50 causes a split:

which in turn causes internal node overflow:
11 causes a split:
finally, with the insertion of 32 ..

we get a split of a leaf..

causing overflow at an internal node, which also splits..
... leading to the final tree

Again, remember that the lowest level of nodes are the leaf nodes. They are linked together left to right (but we don’t draw the links here to avoid clutter).