

Differential Directory (DiDi) – Brief Description

The Differential Directory (DiDi) stores only the difference between each key and the previous one (already sorted), rather than storing full keys in each node.

Much like when writing a list by hand and using quotation marks (“”) to indicate repetition from the line above—for example:

```
John Miles
"    Smith
```

DiDi applies a similar concept, but instead of letters, it identifies the first *differing bit* between keys. It stores this position in a node, which also contains references to the left and right nodes (if they exist), and to the disk record where the full key and its data are stored.

For instance, if the first key inserted is "John Smith", the index is initially empty. The first differing character (compared to “nothing”) is J at position 1, so the index becomes 1J → record 1.

If we then insert "John Miles", we read 1J, confirm it matches at position 1, and read record 1. The index is updated to:

1J → record 1, right → record 2, and then 6S → record 2.

Now, if we insert "John Serrote", we find that position 6 in record 1 is not S, so we follow the pointer to record 2. If it matches, we insert a new node 6e between records 1 and 2, updating the structure to:

1J → record 1, right → record 3,

6S → record 3, right → record 2,

and finally

7m → record 2 (with no children).

With this structure—which is a bit more elaborate—the index remains sorted, and only a maximum of 3 nodes and a single disk read are required per insertion.