Article **Talk**   Read **Edit** **View history**   Search 🔍

Splay tree

From Wikipedia, the free encyclopedia

A **splay tree** is a self-adjusting [binary search tree](#) with the additional property that recently accessed elements are quick to access again. It performs basic operations such as insertion, look-up and removal in O(log n) [amortized](#) time. For many sequences of non-random operations, splay trees perform better than other search trees, even when the specific pattern of the sequence is unknown. The splay tree was invented by [Daniel Dominic Sleator](#) and [Robert Endre Tarjan](#) in 1985.[1]

All normal operations on a binary search tree are combined with one basic operation, called *splaying*. Splaying the tree for a certain element rearranges the tree so that the element is placed at the root of the tree. One way to do this is to first perform a standard binary tree search for the element in question, and then use [tree rotations](#) in a specific fashion to bring the element to the top. Alternatively, a top-down algorithm can combine the search and the tree reorganization into a single phase.

Splay tree		
Type	Tree	
Invented	1985	
Invented by	Daniel Dominic Sleator and Robert Endre Tarjan	
Time complexity in big O notation		
	Average	Worst case
Space	O(n)	O(n)
Search	O(log n)	amortized O(log n)
Insert	O(log n)	amortized O(log n)
Delete	O(log n)	amortized O(log n)

Advantages [edit]

Good performance for a splay tree depends on the fact that it is self-optimizing, in that frequently accessed nodes will move nearer to the root where they can be accessed more quickly. The worst-case height—though unlikely—is O(n), with the average being O(log *n*). Having frequently used nodes near the root is an advantage for many practical applications (also see [Locality of reference](#)), and is particularly useful for implementing [caches](#) and [garbage collection](#) algorithms.

Advantages include:

- Comparable performance—average-case performance is as efficient as other trees.[2]
- Small memory footprint—splay trees do not need to store any bookkeeping data.
- Possibility of creating a [persistent data structure](#) version of splay trees—which allows access to both the previous and new versions after an update. This can be useful in [functional programming](#), and requires amortized O(log *n*) space per update.
- Working well with nodes containing identical keys—contrary to other types of self-balancing trees. Even with identical keys, performance remains amortized O(log *n*). All tree operations preserve the order of the

identical nodes within the tree, which is a property similar to stable sorting algorithms. A carefully designed find operation can return the leftmost or rightmost node of a given key.

Disadvantages [edit]

The most significant disadvantage of splay trees is that the height of a splay tree can be linear. For example, this will be the case after accessing all *n* elements in non-decreasing order. Since the height of a tree corresponds to the worst-case access time, this means that the actual cost of an operation can be high. However the amortized access cost of this worst case is logarithmic, O(log *n*). Also, the expected access cost can be reduced to O(log *n*) by using a randomized variant.[3]

The representation of splay trees can change even when they are accessed in a 'read-only' manner (i.e. by *find* operations). This complicates the use of such splay trees in a multi-threaded environment. Specifically, extra management is needed if multiple threads are allowed to perform *find* operations concurrently. This also makes them unsuitable for general use in purely functional programming, although they can be used in limited ways to implement priority queues even there.

Operations [edit]

Splaying [edit]

When a node *x* is accessed, a splay operation is performed on *x* to move it to the root. To perform a splay operation we carry out a sequence of *splay steps*, each of which moves *x* closer to the root. By performing a splay operation on the node of interest after every access, the recently accessed nodes are kept near the root and the tree remains roughly balanced, so that we achieve the desired amortized time bounds.

Each particular step depends on three factors:

- Whether *x* is the left or right child of its parent node, *p*,
- whether *p* is the root or not, and if not
- whether *p* is the left or right child of *its* parent, *g* (the *grandparent* of x).

It is important to remember to set *gg* (the *great-grandparent* of x) to now point to x after any splay operation. If *gg* is null, then x obviously is now the root and must be updated as such.

There are three types of splay steps, each of which has a left- and right-handed case. For the sake of brevity, only one of these two is shown for each type. These three types are:

Zig step: this step is done when *p* is the root. The tree is rotated on the edge between *x* and *p*. Zig steps exist to deal with the parity issue and will be done only as the last step in a splay operation and only when *x* has odd depth at the beginning of the operation.
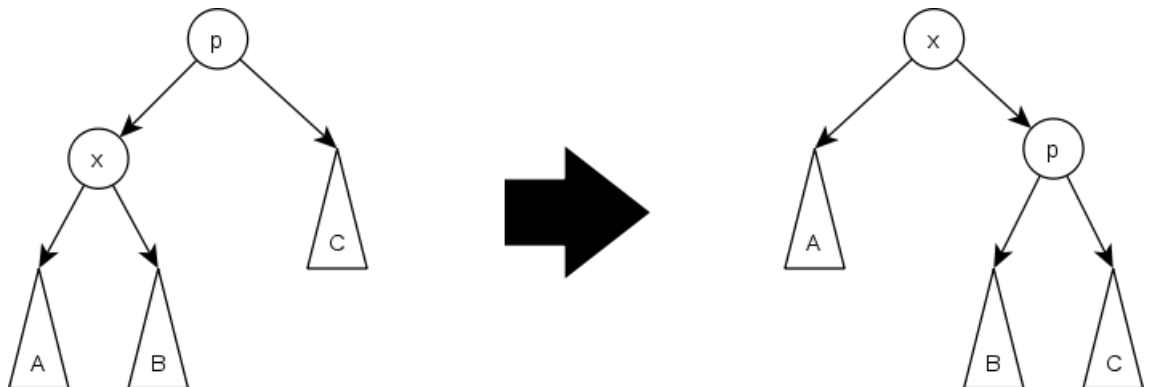
Zig-zig step: this step is done when *p* is not the root and *x* and *p* are either both right children or are both left children. The picture below shows the case where *x* and *p* are both left children. The tree is rotated on the edge joining *p* with *its* parent *g*, then rotated on the edge joining *x* with *p*. Note that zig-zig steps are the only thing that differentiate splay trees from the *rotate to root* method introduced by Allen and Munro[4] prior to the introduction of splay trees.
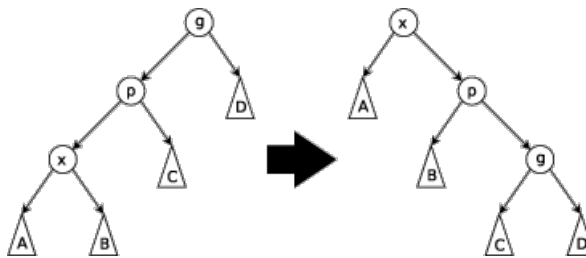
Zig-zag step: this step is done when *p* is not the root and *x* is a right child and *p* is a left child or vice versa. The tree is rotated on the edge between *p* and x, and then rotated on the resulting edge between *x* and g.
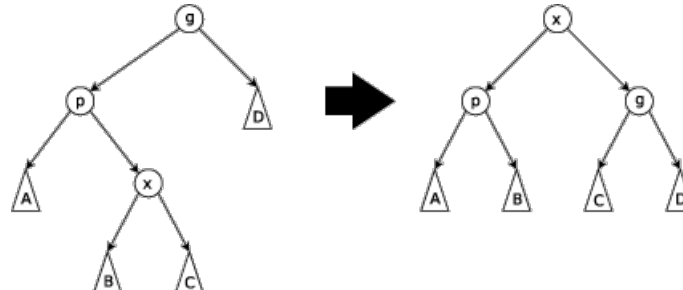
Join   [edit]

Given two trees S and T such that all elements of S are smaller than the elements of T, the following steps can be used to join them to a single tree:

- Splay the largest item in S. Now this item is in the root of S and has a null right child.
- Set the right child of the new root to T.

Split   [edit]

Given a tree and an element *x*, return two new trees: one containing all elements less than or equal to *x* and the other containing all elements greater than *x*. This can be done in the following way:

- Splay *x*. Now it is in the root so the tree to its left contains all elements smaller than *x* and the tree to its right contains all element larger than *x*.
- Split the right subtree from the rest of the tree.

Insertion   [edit]

To insert a value *x* into a splay tree:

- Insert *x* as with a normal binary search tree.
- Splay the newly inserted node *x* to the top of the tree.

ALTERNATIVE:

- Use the split operation to split the tree at the value of *x* to two sub-trees: S and T.
- Create a new tree in which *x* is the root, S is its left sub-tree and T its right sub-tree.

Deletion   [edit]

To delete a node *x*, use the same method as with a binary search tree: if *x* has two children, swap its value with that of either the rightmost node of its left sub tree (its in-order predecessor) or the leftmost node of its right subtree (its in-order successor). Then remove that node instead. In this way, deletion is reduced to the problem of removing a node with 0 or 1 children. Unlike a binary search tree, in a splay tree after deletion, we splay the parent of the removed node to the top of the tree.

ALTERNATIVE:

- The node to be deleted is first splayed, i.e. brought to the root of the tree and then deleted. leaves the tree with two sub trees.
- The two sub-trees are then joined using a "join" operation.

Implementation and variants   [edit]

Splaying, as mentioned above, is performed during a second, bottom-up pass over the access path of a node. It is possible to record the access path during the first pass for use during the second, but that requires extra space during the access operation. Another alternative is to keep a parent pointer in every node, which avoids the need for extra space during access operations but may reduce overall time efficiency because of the need

to update those pointers.[1]

Another method which can be used is based on the argument that we can restructure the tree on our way down the access path instead of making a second pass. This top-down splaying routine uses three sets of nodes - left tree, right tree and middle tree. The first two contain all items of original tree known to be less than or greater than current item respectively. The middle tree consists of the sub-tree rooted at the current node. These three sets are updated down the access path while keeping the splay operations in check. Another method, semisplaying, modifies the zig-zig case to reduce the amount of restructuring done in all operations. [1][5]

Below there is an implementation of splay trees in C++, which uses pointers to represent each node on the tree. This implementation is based on bottom-up splaying version and uses the second method of deletion on a splay tree. Also, unlike the above definition, this C++ version does *not* splay the tree on finds - it only splays on insertions and deletions.

```cpp
#include <functional>

#ifndef SPLAY_TREE
#define SPLAY_TREE

template< typename T, typename Comp = std::less< T > >
class splay_tree {
private:
  Comp comp;
  unsigned long p_size;

  struct node {
    node *left, *right;
    node *parent;
    T key;
    node( const T& init = T( ) ) : left( 0 ), right( 0 ), parent( 0 ), key( init ) {
}
    ~node( ) {
      if( left ) delete left;
      if( right ) delete right;
      if( parent ) delete parent;
    }
  } *root;

  void left_rotate( node *x ) {
    node *y = x->right;
    if(y) {
      x->right = y->left;
      if( y->left ) y->left->parent = x;
      y->parent = x->parent;
    }

    if( !x->parent ) root = y;
    else if( x == x->parent->left ) x->parent->left = y;
    else x->parent->right = y;
    if(y) y->left = x;
    x->parent = y;
  }

  void right_rotate( node *x ) {
    node *y = x->left;
    if(y) {
      x->left = y->right;
      if( y->right ) y->right->parent = x;
      y->parent = x->parent;
    }
    if( !x->parent ) root = y;
    else if( x == x->parent->left ) x->parent->left = y;
    else x->parent->right = y;
    if(y) y->right = x;
    x->parent = y;
  }

  void splay( node *x ) {
    while( x->parent ) {
```

```cpp
        if( !x->parent->parent ) {
          if( x->parent->left == x ) right_rotate( x->parent );
          else left_rotate( x->parent );
        } else if( x->parent->left == x && x->parent->parent->left == x->parent ) {
          right_rotate( x->parent->parent );
          right_rotate( x->parent );
        } else if( x->parent->right == x && x->parent->parent->right == x->parent ) {
          left_rotate( x->parent->parent );
          left_rotate( x->parent );
        } else if( x->parent->left == x && x->parent->parent->right == x->parent ) {
          right_rotate( x->parent );
          left_rotate( x->parent );
        } else {
          left_rotate( x->parent );
          right_rotate( x->parent );
        }
      }
    }

    void replace( node *u, node *v ) {
      if( !u->parent ) root = v;
      else if( u == u->parent->left ) u->parent->left = v;
      else u->parent->right = v;
      if( v ) v->parent = u->parent;
    }

    node* subtree_minimum( node *u ) {
      while( u->left ) u = u->left;
      return u;
    }

    node* subtree_maximum( node *u ) {
      while( u->right ) u = u->right;
      return u;
    }
  public:
    splay_tree( ) : root( 0 ), p_size( 0 ) { }

    void insert( const T &key ) {
      node *z = root;
      node *p = 0;

      while( z ) {
        p = z;
        if( comp( z->key, key ) ) z = z->right;
        else z = z->left;
      }

      z = new node( key );
      z->parent = p;

      if( !p ) root = z;
      else if( comp( p->key, z->key ) ) p->right = z;
      else p->left = z;

      splay( z );
      p_size++;
    }

    node* find( const T &key ) {
      node *z = root;
      while( z ) {
        if( comp( z->key, key ) ) z = z->right;
        else if( comp( key, z->key ) ) z = z->left;
        else return z;
      }
      return 0;
    }

    void erase( const T &key ) {
      node *z = find( key );
      if( !z ) return;
```

```
      splay( z );

      if( !z->left ) replace( z, z->right );
      else if( !z->right ) replace( z, z->left );
      else {
        node *y = subtree_minimum( z->right );
        if( y->parent != z ) {
          replace( y, y->right );
          y->right = z->right;
          y->right->parent = y;
        }
        replace( z, y );
        y->left = z->left;
        y->left->parent = y;
      }

      delete z;
      p_size--;
    }
  }

  const T& minimum( ) { return subtree_minimum( root )->key; }
  const T& maximum( ) { return subtree_maximum( root )->key; }

  bool empty( ) const { return root == 0; }
  unsigned long size( ) const { return p_size; }
};

#endif // SPLAY_TREE
```

Analysis [edit]

A simple amortized analysis of static splay trees can be carried out using the potential method. Define:

- size(r) - the number of nodes in the sub-tree rooted at node r (including r).
- rank(r) = $\log_2$(size(r)).
- Φ = the sum of the ranks of all the nodes in the tree.

Φ will tend to be high for poorly balanced trees and low for well-balanced trees.

To apply the potential method, we first calculate $\Delta\Phi$ - the change in the potential caused by a splay operation. We check each case separately. Denote by rank' the rank function after the operation. x, p and g are the nodes affected by the rotation operation (see figures above).

Zig step:

$\Delta\Phi$ = rank'(p) - rank(p) + rank'(x) - rank(x) [since only p and x change ranks]

= rank'(p) - rank(x) [since rank'(x)=rank(p)]

≤ rank'(x) - rank(x) [since rank'(p)<rank'(x)]

Zig-Zig step:

$\Delta\Phi$ = rank'(g) - rank(g) + rank'(p) - rank(p) + rank'(x) - rank(x)

≤ rank'(g) + rank'(p) - rank(p) - rank(x) [since rank'(x)=rank(g)]

≤ rank'(g) + rank'(x) - 2 rank(x) [since rank(x)<rank(p) and rank'(x)>rank'(p)]

≤ 3(rank'(x)-rank(x)) - 2 [due to the convexity of the log function]

Zig-Zag step:

$\Delta\Phi$ = rank'(g) - rank(g) + rank'(p) - rank(p) + rank'(x) - rank(x)

≤ rank'(g) + rank'(p) - 2 rank(x) [since rank'(x)=rank(g) and rank(x)<rank(p)]

≤ 2(rank'(x)-rank(x)) - 2 [due to the convexity of the log function]

The amortized cost of any operation is $\Delta\Phi$ plus the actual cost. The actual cost of any zig-zig or zig-zag operation is 2 since there are two rotations to make. Hence:

amortized-cost = cost + $\Delta\Phi$

≤ 3(rank'(x)-rank(x))

When summed over the entire splay operation, this telescopes to 3(rank(root)-rank(x)) which is O(log n). The Zig operation adds an amortized cost of 1, but there's at most one such operation.

So now we know that the total *amortized* time for a sequence of *m* operations is:

$$T_{\text{amortized}}(m) = O(m \log n)$$

To go from the amortized time to the actual time, we must add the decrease in potential from the initial state before any operation is done (Φ_i) to the final state after all operations are completed (Φ_f).

$$\Phi_i - \Phi_f = \sum_x \text{rank}_i(x) - \text{rank}_f(x) = O(n \log n)$$

where the last inequality comes from the fact that for every node x, the minimum rank is 0 and the maximum rank is log(n).

Now we can finally bound the actual time:

$$T_{\text{actual}}(m) = O(m \log n + n \log n)$$

Weighted analysis   [edit]

The above analysis can be generalized in the following way.

- Assign to each node r a weight $w(r)$.
- Define size(r) = the sum of weights of nodes in the sub-tree rooted at node r (including r).
- Define rank(r) and Φ exactly as above.

The same analysis applies and the amortized cost of a splaying operation is again:

$$\text{rank}(root) - \text{rank}(x) = O(\log W - \log w(x)) = O\left(\log \frac{W}{w(x)}\right)$$

where W is the sum of all weights.

The decrease from the initial to the final potential is bounded by:

$$\Phi_i - \Phi_f \le \sum_{x \in tree} \log \frac{W}{w(x)}$$

since the maximum size of any single node is W and the minimum is $w(x)$.

Hence the actual time is bounded by:

$$O\left(\sum_{x \in sequence} \left(\log \frac{W}{w(x)}\right) + \sum_{x \in tree} \left(\log \frac{W}{w(x)}\right) \right)$$

Performance theorems   [edit]

There are several theorems and conjectures regarding the worst-case runtime for performing a sequence S of m accesses in a splay tree containing n elements.

Balance Theorem

The cost of performing the sequence S is $O\left[m \log n + n \log n\right]$ (**Proof**: take a constant weight, e.g. $w(x)$=1 for every node x. Then W=n). This theorem implies that splay trees perform as well as static balanced binary search trees on sequences of at least n accesses.[1]

Static Optimality Theorem

Let q_x be the number of times element x is accessed in S. The cost of performing S is $O\left[m + \sum_{x \in tree} q_x \log \frac{m}{q_x}\right]$ (**Proof**: let $w(x) = q_x$. Then $W = m$). This theorem implies that splay trees perform as well as an optimum static binary search tree on sequences of at least n accesses. They spend less time on the more frequent items.[1]

Static Finger Theorem

Assume that the items are numbered from 1 through n in ascending order. Let f be any fixed element (the 'finger'). Then the cost of performing S is $O\left[m + n \log n + \sum_{x \in sequence} \log(|x - f| + 1)\right]$ (**Proof**: let $w(x) = 1/(|x - f| + 1)^2$. Then W=O(1). The net potential drop is O ($n \log n$) since the weight of any item is at least 1/n^2).[1]

Dynamic Finger Theorem

Assume that the 'finger' for each step accessing an element y is the element accessed in the previous step, x. The cost of performing S is $O\left[m + n + \sum_{x,y \in sequence}^{m} \log(|y - x| + 1)\right]$.[6][7]

Working Set Theorem

At any time during the sequence, let $t(x)$ be the number of distinct elements accessed before the previous time element x was accessed. The cost of performing S is

$$O\left[m + n\log n + \sum_{x \in sequence} \log(t(x)+1)\right]$$ (**Proof**: let $w(x) = 1/(t(x)+1)^2$. Note that

here the weights change during the sequence. However, the sequence of weights is still a permutation of 1, 1/4, 1/9, ..., 1/n^2. So as before W=O(1). The net potential drop is O ($n \log n$).). This theorem is equivalent to splay trees having key-independent optimality.[1]

Scanning Theorem

Also known as the **Sequential Access Theorem** or the **Queue theorem**. Accessing the n elements of a splay tree in symmetric order takes O(n) time, regardless of the initial structure of the splay tree.[8] The tightest upper bound proven so far is $4.5n$.[9]

Dynamic optimality conjecture [edit]

Main article: Optimal binary search tree

In addition to the proven performance guarantees for splay trees there is an unproven conjecture of great interest from the original Sleator and Tarjan paper. This conjecture is known as the *dynamic optimality conjecture* and it basically claims that splay trees perform as well as any other binary search tree algorithm up to a constant factor.

List of unsolved problems in computer science

Do splay trees perform as well as any other binary search tree algorithm?

Dynamic Optimality Conjecture:[1] Let A be any binary search tree algorithm that accesses an element x by traversing the path from the root to x at a cost of $d(x) + 1$, and that between accesses can make any rotations in the tree at a cost of 1 per rotation. Let $A(S)$ be the cost for A to perform the sequence S of accesses. Then the cost for a splay tree to perform the same accesses is $O[n + A(S)]$.

There are several corollaries of the dynamic optimality conjecture that remain unproven:

Traversal Conjecture:[1] Let T_1 and T_2 be two splay trees containing the same elements. Let S be the sequence obtained by visiting the elements in T_2 in preorder (i.e., depth first search order). The total cost of performing the sequence S of accesses on T_1 is $O(n)$.

Deque Conjecture:[8][10][11] Let S be a sequence of m double-ended queue operations (push, pop, inject, eject). Then the cost of performing S on a splay tree is $O(m + n)$.

Split Conjecture:[5] Let S be any permutation of the elements of the splay tree. Then the cost of deleting the elements in the order S is $O(n)$.

Variants [edit]

In order to reduce the number of restructuring operations, it is possible to replace the splaying with *semi-splaying*, in which an element is splayed only halfway towards the root.[1]

Another way to reduce restructuring is to do full splaying, but only in some of the access operations - only when the access path is longer than a threshold, or only in the first m access operations.[1]

See also [edit]

- Finger tree
- Link/cut tree
- Scapegoat tree
- Zipper (data structure)
- Trees
- Tree rotation
- AVL tree
- B-tree
- T-tree
- List of data structures
- Iacono's working set structure
- Geometry of binary search trees

- Splaysort, a sorting algorithm using splay trees

Notes [edit]

1. ^ *a b c d e f g h i j k* Sleator, Daniel D.; Tarjan, Robert E. (1985), "Self-Adjusting Binary Search Trees" 🗎 (PDF), *Journal of the ACM (Association for Computing Machinery)* **32** (3): 652–686, doi:10.1145/3828.3835 🖉
2. ^ Goodrich, Michael; Tamassia, Roberto; Goldwasser, Michael. *Data Structures and Algorithms in Java* (6 ed.). John Wiley & Sons, Inc. p. 506. ISBN 978-1-118-77133-4. "The surprising thing about splaying is that it allows us to guarantee a logarithmic amortized running time, for insertions, deletions, and searches."
3. ^ "Randomized Splay Trees: Theoretical and Experimental Results" 🗎 (PDF). Retrieved 31 May 2011.
4. ^ Allen, Brian; and Munro, Ian (1978), "Self-organizing search trees", *Journal of the ACM* **25** (4): 526–535, doi:10.1145/322092.322094 🖉
5. ^ *a b* Lucas, Joan M. (1991), "On the Competitiveness of Splay Trees: Relations to the Union-Find Problem", *Online Algorithms, Center for Discrete Mathematics and Theoretical Computer Science (DIMACS) Series in Discrete Mathematics and Theoretical Computer Science Vol. 7*: 95–124
6. ^ Cole, Richard; Mishra, Bud; Schmidt, Jeanette; and Siegel, Alan (2000), "On the Dynamic Finger Conjecture for Splay Trees. Part I: Splay Sorting log n-Block Sequences", *SIAM Journal on Computing* **30**: 1–43, doi:10.1137/s0097539797326988 🖉
7. ^ Cole, Richard (2000), "On the Dynamic Finger Conjecture for Splay Trees. Part II: The Proof", *SIAM Journal on Computing* **30**: 44–85, doi:10.1137/S009753979732699X 🖉
8. ^ *a b* Tarjan, Robert E. (1985), "Sequential access in splay trees takes linear time", *Combinatorica* **5** (4): 367–378, doi:10.1007/BF02579253 🖉
9. ^ Elmasry, Amr (2004), "On the sequential access theorem and Deque conjecture for splay trees", *Theoretical Computer Science* **314** (3): 459–466, doi:10.1016/j.tcs.2004.01.019 🖉
10. ^ Pettie, Seth (2008), "Splay Trees, Davenport-Schinzel Sequences, and the Deque Conjecture", *Proceedings of the 19th ACM-SIAM Symposium on Discrete Algorithms* **0707**: 1115–1124, arXiv:0707.2160 🖉, Bibcode:2007arXiv0707.2160P 🖉
11. ^ Sundar, Rajamani (1992), "On the Deque conjecture for the splay algorithm", *Combinatorica* **12** (1): 95–124, doi:10.1007/BF01191208 🖉

References [edit]

- Knuth, Donald. *The Art of Computer Programming*, Volume 3: *Sorting and Searching*, Third Edition. Addison-Wesley, 1997. ISBN 0-201-89685-0. Page 478 of section 6.2.3.

External links [edit]

- NIST's Dictionary of Algorithms and Data Structures: Splay Tree 🖉
- Implementations in C and Java (by Daniel Sleator) 🖉
- Pointers to splay tree visualizations 🖉
- Fast and efficient implentation of Splay trees 🖉
- Top-Down Splay Tree Java implementation 🖉
- Zipper Trees 🖉
- splay tree video 🖉

v · t · e	Tree data structures	[show]
v · t · e	Data structures	[show]

Categories: Binary trees | Search trees

WIKIMEDIA project   Powered By MediaWiki