

1 Pseudocode

Algorithm Tree Fields Description

◇ Shared

- *Tree* *tree* : A binary tree of Nodes. *root* is a pointer to the root node.

◇ Local

- **Node* *leaf* : a pointer to the process's leaf in the tree.

◇ Structures

► Node

- **Node* *left*, *right*, *parent* : initialized when creating the tree.
- *BlockList* *blocks* implemented with an array.
- *int* *size*= 1: #blocks in blocks.
- *int* *num*_{propagated}= 0 : # groups of blocks that have been propagated from the node to its parent. Since it is incremented after propagating, it may be behind by 1.
- *int*[] *super*: *super*[*i*] stores an approximate index of the superblock of the blocks in *blocks* whose group field have value *i*.

► *Block* ▷ For a block in a blocklist we define *the prefix for the block* to be the blocks in the *BlockList* up to and including the block.
put the definitions before the pseudocode

- *int* *group* : the value read from *num*_{propagated} when appending this block to the node.

► LeafBlock extends Block

- *Object* *element* : Each block in a leaf represents a single operation. For enqueue operations *element* is the input of the enqueue and for dequeue operations it is null.
- *Object* *response* : stores the response of the operation in the *LeafBlock*.
- *int* *sum*_{enq}, *sum*_{deq} : # enqueue, dequeue operations in the prefix for the block

► InternalBlock extends Block

- *int* *end*_{left}, *end*_{right} : index of the last subblock of the block in the left and right child
- *int* *sum*_{enq-left} : # enqueue operations in the prefix for *left.blocks[end*_{left}]
- *int* *sum*_{deq-left} : # dequeue operations in the prefix for *left.blocks[end*_{left}]
- *int* *sum*_{enq-right} : # enqueue operations in the prefix for *right.blocks[end*_{right}]
- *int* *sum*_{deq-right} : # dequeue operations in the prefix for *right.blocks[end*_{right}]

► RootBlock extends InternalBlock

- *int* *length* : length of the queue after performing all operations in the prefix for this block
-

Algorithm BlockList

▷ : Supports two operations *blocks.tryAppend(Block b)*, *blocks[i]*. Initially empty, when *blocks.tryAppend(b, n)*

returns true *b* is appended to *blocks[n]* and *blocks[i]* returns *i*th block in the blocks. If some instance of *blocks.tryAppend(b, n)* returns false there is a concurrent instance of *blocks.tryAppend(b', n)* which has returned true. *blocks[0]* contains an empty block with all fields equal to 0 and *end*_{left}, *end*_{right} pointers to the first block of the corresponding children.

◇ Array implementation

blocks[]): array of blocks

701: *boolean* TRYAPPEND(*block* *blk*, *int* *n*)

702: *return* CAS(*blocks*[*n*], null, *blk*)

703: *end* TRYAPPEND

Variable naming:

- *b*_{op}: index of the block containing operation *op*
- *r*_{op}: rank of operation *op* i.e. the ordering among the operations of its type according to linearization ordering

Abbreviations:

- *blocks*[*b*].*sum*_{*x*}=*blocks*[*b*].*sum*_{*x-left*}+*blocks*[*b*].*sum*_{*x-right*} (for *b*≥0 and *x* ∈ {enq, deq})
- *blocks*[*b*].*sum*=*blocks*[*b*].*sum*_{enq}+*blocks*[*b*].*sum*_{deq} (for *b*≥0)
- *blocks*[*b*].*num*_{*x*}=*blocks*[*b*].*sum*_{*x*}-*blocks*[*b-1*].*sum*_{*x*}
(for *b*>0 and *x* ∈ {∅, enq, deq, enq-left, enq-right, deq-left, deq-right}, *blocks*[0].*num*_{*x*}=0)

Algorithm Queue

```
201: void ENQUEUE(Object e) ▷ Creates a block with element e and appends
    it to the tree.
202:   block newBlock= NEW(LeafBlock)
203:   newBlock.element= e
204:   newBlock.sumenq= leaf.blocks[leaf.size].sumenq+1
205:   newBlock.sumdeq= leaf.blocks[leaf.size].sumdeq
206:   leaf.APPEND(newBlock)
207: end ENQUEUE

208: <int, int> FINDRESPONSE(int b, int i) ▷ Computes the rank and
    index of the block in the root of the enqueue that is the response of the ith
    dequeue in the root's bth block. Returns <-1,--> if the queue is empty.
209:   if root.blocks[b-1].length + root.blocks[b].numenq - i < 0 then
210:     return <-1,-->
211:   else
    ▷ We call the dequeues that
    return a value non-null dequeues. rth non-null dequeue returns the element
    of th rth enqueue. We can compute # non-null dequeues in the prefix for
    a block this way: #non-null dequeues= length - #enqueues. Note that the
    ith dequeue in the given block is not a non-null dequeue.
212:     renq= root.blocks[b-1].sumenq- root.blocks[b-1].length + i
213:     return <root.DSEARCH(renq, b), renq>
214:   end if
215: end FINDRESPONSE

216: Object DEQUEUE()
217:   block newBlock= NEW(LeafBlock) ▷ Creates a block
    with null value element, appends it to the tree, computes its order among
    operations, then computes and returns its response.
218:   newBlock.element= null
219:   newBlock.sumenq= leaf.blocks[leaf.size].sumenq
220:   newBlock.sumdeq= leaf.blocks[leaf.size].sumdeq+1
221:   leaf.APPEND(newBlock)
222:   <bdeq, rdeq>= INDEXDEQ(leaf.size, 1) ▷ r is the rank among the
    dequeues of the dequeue of the bdeqth block in the root containing.
223:   <benq, renq>= FINDRESPONSE(bdeq, rdeq) ▷ renq is the rank
    of the enqueue whose element is the response to the dequeue in the block
    containing it and bdeq is the index of that block of it in the blocklist. If the
    response is null then rdeq is -1.
224:   if renq== -1 then
225:     output= null
226:   else
227:     output= GETENQ(benq, renq) ▷ getting the reponse's element.
228:   end if
229:   return output
230: end DEQUEUE
```

deqRest

Algorithm Node

```

301: void PROPAGATE()
302:   if not REFRESH() then
303:     REFRESH()
304:   end if
305:   if this is not root then
306:     parent.PROPAGATE()
307:   end if
308: end PROPAGATE

309: boolean REFRESH()
310:   s = size
311:   <new, npleft, npright> = CREATEBLOCK(s)
312:   if new.num==0 then return true
313:   else if blocks.tryAppend(new, s) then
314:     for each dir in {left, right} do
315:       CAS(dir.super[npdir], null, h+1)
316:       CAS(dir.numpropagated, npdir, npdir+1)
317:     end for
318:     CAS(size, s, s+1)
319:     return true
320:   else
321:     CAS(size, s, s+1)
322:     return false
323:   end if
324: end REFRESH

325: int BSEARCH(field f, int i, int start, int end)
326: end BSEARCH

327: int DSEARCH(int i, int end)
328:   start = b-1
329:   while root.blocks[start].sumenq ≥ renq do
330:     start = start - (b - start)
331:   end while return root.BSearch(sumenq, renq, start, b)
332: end DSEARCH

333: <Block, int, int> CREATEBLOCK(int i)
334:   block newBlock = NEW(block)
335:   newBlock.group = numpropagated
336:   newBlock.order = i
337:   for each dir in {left, right} do
338:     indexlast = dir.size
339:     indexprev = blocks[i-1].enddir
340:     newBlock.enddir = indexlast
341:     blocklast = dir.blocks[indexlast]
342:     blockprev = dir.blocks[indexprev]
343:     newBlock.sumenq-dir = blocks[i-1].sumenq-dir + blocklast.sumenq
344:     npdir = dir.numpropagated
345:     newBlock.sumdeq-dir = blocks[i-1].sumdeq-dir + blocklast.sumdeq
346:   end for
347:   if this is root then
348:     newBlock.length = max(root.blocks[i-1].length + b.numenq - b.numdeq, 0)
349:   end if
350:   return <b, npleft, npright>
351: end CREATEBLOCK

```

firstRefresh: 302: if not REFRESH() then
 secondRefresh: 303: REFRESH()
 304: end if
 305: if this is not root then
 306: parent.PROPAGATE()
 307: end if
 308: end PROPAGATE

309: boolean REFRESH()
 readSize: 310: s = size
 311: <new, np_{left}, np_{right}> = CREATEBLOCK(s)
 312: if new.num==0 then return true
 313: else if blocks.tryAppend(new, s) then
 okcas: 314: for each dir in {left, right} do
 315: CAS(dir.super[np_{dir}], null, h+1)
 316: CAS(dir.num_{propagated}, np_{dir}, np_{dir}+1)
 317: end for
 incrementHead1: 318: CAS(size, s, s+1)
 319: return true
 320: else
 321: CAS(size, s, s+1)
 incrementHead2: 322: return false
 323: end if
 324: end REFRESH

325: int BSEARCH(field f, int i, int start, int end)
 326: end BSEARCH

327: int DSEARCH(int i, int end)
 328: start = b-1
 329: while root.blocks[start].sum_{enq} ≥ r_{enq} do
 330: start = start - (b - start)
 331: end while return root.BSearch(sum_{enq}, r_{enq}, start, b)
 332: end DSEARCH

333: <Block, int, int> CREATEBLOCK(int i)
 334: block newBlock = NEW(block)
 335: newBlock.group = num_{propagated}
 336: newBlock.order = i
 337: for each dir in {left, right} do
 338: index_{last} = dir.size
 339: index_{prev} = blocks[i-1].end_{dir}
 340: newBlock.end_{dir} = index_{last}
 341: block_{last} = dir.blocks[index_{last}]
 342: block_{prev} = dir.blocks[index_{prev}]
 343: newBlock.sum_{enq-dir} = blocks[i-1].sum_{enq-dir} + block_{last}.sum_{enq}
 344: np_{dir} = dir.num_{propagated}
 345: newBlock.sum_{deq-dir} = blocks[i-1].sum_{deq-dir} + block_{last}.sum_{deq}
 346: end for
 347: if this is root then
 348: newBlock.length = max(root.blocks[i-1].length + b.num_{enq} - b.num_{deq}, 0)
 349: end if
 350: return <b, np_{left}, np_{right}>
 351: end CREATEBLOCK

Algorithm Node

↪ Precondition: $\text{blocks}[b].\text{num}_{\text{enq}} \geq i$

```
401: element GETENQ(int b, int i)
402:   if this is leaf then
403:     return blocks[b].element
404:   else if  $i \leq \text{blocks}[b].\text{num}_{\text{enq-left}}$  then                                ▷ i exists in the left child of this node
405:     subBlock= left.BSEARCH( $\text{sum}_{\text{enq}}$ , i,  $\text{blocks}[b-1].\text{end}_{\text{left}}+1$ ,  $\text{blocks}[b].\text{end}_{\text{left}}$ )    ▷ Search range of left child's subblocks of blocks[b].
406:     return left.GET( $i-\text{left.blocks}[\text{subBlock}-1].\text{sum}_{\text{enq}}$ , subBlock)
407:   else
408:     i=  $i-\text{blocks}[b].\text{num}_{\text{enq-left}}$ 
409:     subBlock= right.BSEARCH( $\text{sum}_{\text{enq}}$ , i,  $\text{blocks}[b-1].\text{end}_{\text{right}}+1$ ,  $\text{blocks}[b].\text{end}_{\text{right}}$ )    ▷ Search range of right child's subblocks of blocks[b].
410:     return right.GET( $i-\text{right.blocks}[\text{subBlock}-1].\text{sum}_{\text{enq}}$ , subBlock)
411:   end if
412: end GETENQ
```

↪ Precondition: b th block of the node has propagated up to the root and $\text{blocks}[b].\text{num}_{\text{enq}} \geq i$.

```
413: <int, int> INDEXDEQ(int b, int i)                                ▷ Returns the rank of  $i$ th dequeue in the  $b$ th block of the node, among the dequeues in the root.
414:   if this is root then
415:     return <b, i>
416:   else
417:     dir= (parent.left==n)? left: right                                ▷ check if a left or a right child
418:     superBlock= parent.BSEARCH( $\text{sum}_{\text{deq-dir}}$ , i,  $\text{super}[\text{blocks}[b].\text{group}]-p$ ,  $\text{super}[\text{blocks}[b].\text{group}]+p$ )
                                                                ▷ superblock's group has at most  $p$  difference with the value stored in super[].
419:     if dir is right then
420:       i+=  $\text{blocks}[\text{superBlock}].\text{sum}_{\text{deq-left}}$                                 ▷ consider the dequeues from the right child
421:     end if
422:     return this.parent.INDEXDEQ(superBlock, i)
423:   end if
424: end INDEX
```

Algorithm Leaf

```
601: void APPEND(block blk)                                ▷ Append is only called by the owner of the leaf.
602:   size+=1
603:   blk.group= size
604:   blocks[size]= blk
605:   parent.PROPAGATE()
606: end APPEND
```

2 Proof of Linearizability

TEST As a temporary test I have changed the name of `n.size` to `n.head` here, other options are `n.head` and `n.lastBlock` but they might be confusing since we have used them before. Fix the logical order of definitions (cyclic references).

TODO Fallback safety lemmas. Some parts are obsolete.

Questions When I write the lemmas since every claim in my mind is correct maybe I miss some fact that need proof or maybe I refer to some lemmas that are generally correct but not needed for the linearizability proof. Is lemma 7 necessary? Is lemma 13 induced trivially from lemma 8?

TEST Is it better to show $\text{ops}(\text{EST}_{n, t})$ with $\text{EST}_{n, t}$?

TEST How to merge notions of blocks and operations? block `b` \sqsubseteq block `c` means `b` is subblock of `c`. block `b` \in set `B` means `b` is in `B`. Merge these two to have shorter formulaes.

Definition 1 (Block). A block is an object storing some statistics, as described in Algorithm Queue. It implicitly represents a set of operations. If `n.blocks[i]==b` we call `i` the *index* of block `b`. Block `b` is before block `b'` in node `n` if and only if `b`'s index is smaller than `b'`'s.

Definition 2 (Subblock). Block `b` is a *direct subblock* of `n.blocks[i]` if it is $\in \text{n.left.blocks}[\text{n.blocks[i-1].end_left+1} \dots \text{n.blocks[i].end_left}] \cup \text{n.right.blocks}[\text{n.blocks[i-1].end_right+1} \dots \text{n.blocks[i].end_right}]$. Block `b` is a subblock of a `n.blocks[i]` if it is a direct subblock of it or subblock of a direct subblock of it. Block `b` is direct superblock of block `c` if `c` is direct subblock of `b`. The set of operations of block `b` are the operations in the leaf sub blocks of `b`. We show the set of operations of block by $\text{ops}(b)$. We also say `b` contains *op* if $op \in \text{ops}(b)$.

For simplicity we say block `b` is propagated to node `n` or to a set of blocks `S` if `b` is in `n.blocks` or `S` or is a subblock of a block in `n.blocks` or `S`.

Definition 3. A block `b` in `n.blocks` is *established* at time `t` if `n.head` is greater than the index of `b` in `n.blocks` at time `t`. Block `b` is in $\text{EST}_{n, t}$ if `b` is a subblock of `b'` in `n.blocks` such that `b'` is established at time `t`.

head **Observation 4.** Once a block `b` is written in `n.blocks[i]` then `n.blocks[i]` never changes.

headProgress **Lemma 5** (headProgress). `n.head` is non-decreasing over time and $\text{n.blocks[i].end_left} \geq \text{n.blocks[i-1].end_left}, \text{n.blocks[i].end_right} \geq \text{n.blocks[i-1].end_right}$.

Proof. The first claim follows trivially from the pseudocode since `n.head` is only incremented. Also when `n.blocks[i]` is created its `end_left`, `end_right` are greater than or equal to the values in `n.blocks[i-1]`. Since $\text{blocks[i-1].end_dir} < \text{dir.size} = \text{blocks[i].end_dir}$ (Lines lastLine, prevLine `l??`). □

Lemma 6. Every block has most one direct superblock.

Proof. To show this we are going to refer to the way `n.blocks[]` is partitioned while propagating blocks up to `n.parent`. `n.CreateBlock(i)` merges the blocks in $\text{n.left.blocks}[\text{n.blocks[i-1].end_left} \dots \text{n.blocks[i].end_left}]$ and $\text{n.right.blocks}[\text{n.blocks[i-1].end_right} \dots \text{n.blocks[i].end_right}]$ (Lines lastLine, pr `l??`). Since `end_left`, `end_right` are non-decreasing, so the range of the subblocks of `n.blocks[i]` which is $(\text{n.blocks[i-1].end_dir}+1 \dots \text{n.blocks[i].end_dir})$ does not overlap with the range of the subblocks of `n.blocks[i-1]`. □

append **Corollary 7** (No Duplicates). If *op* is in `n.blocks[i]` then there is no `j` $\neq i$ such that $op \in \text{ops}(\text{n.blocks[j]})$.

headPosition **Invariant 8** (headPosition). If the value of `n.head` is `h` then, `n.blocks[i]=null` for `i>h` and `n.blocks[i]≠null` for `i<h`.

Proof. The invariant is true initially since 1 is assigned to `n.head` and `n.blocks[x]` is null for every `x`. The truth of the invariant may be affected by writing into `n.blocks` or incrementing `n.head`.

Some value is written into `n.blocks[head]` only in Line 313. It is obvious that writing into `n.blocks[head]` preserves the invariant. The value of `n.head` is modified only in lines ^{incrementHead1}318, ^{incrementHead2}321. Depending on whether the `TryAppend()` in Line ^{cas}313 succeeded or not we show that the claim holds after the increment lines of `n.head` in either case. If `head` is incremented to h it is sufficient to show `n.blocks[h] ≠ null` to prove the invariant still holds. In the first case the process applied a successful `TryAppend(new, h)` in line ^{okcas}314, which means `n.blocks[h]` is not null anymore. Note that whether ^{incrementHead1}318 returns true or false after Line `n.head` we know has been incremented from Line ^{readSize}310. The failure case is also the same since it means some value is written into `n.blocks[head]` by some process. \square

Explain More

Lemma 9 (^{establishedOrder}establishedOrder). *If time $t < \text{time } t'$, then $\text{ops}(\text{EST}_n, t) \subseteq \text{ops}(\text{EST}_n, t')$.*

Proof. Blocks are only appended (not modified) with CAS to `n.blocks[n.head]` and `n.head` is non-decreasing, so the set of operations in established blocks of a node can only grow. \square

`CreateBlock()` aggregates the blocks in the children that are not already established in the parent into one block. If a `Refresh()` procedure returns true it means it has appended the block created by `CreateBlock()` into the parent node's sequence. So suppose two `Refreshes` fail. Since the first `Refresh()` was not successful, it means another CAS operation by a `Refresh`, concurrent to the first `Refresh()`, was successful before the second `Refresh()`. So it means the second failed `Refresh` is concurrent with another successful `Refresh()` that assuredly has read block before the mentioned line 35. After all it means if any of the `Refresh()` attempts were successful the claim is true, and also if both fail the mentioned claim still holds.

Lemma 10 (^{trueRefresh}trueRefresh). *Let t_i be the time a `n.Refresh()` operation is invoked and t_t be the time it terminates. Suppose `n.Refresh()`'s `TryAppend(new, s)` returns true, then $\text{ops}(\text{EST}_{n.\text{left}}, t_i) \cup \text{ops}(\text{EST}_{n.\text{right}}, t_i) \subseteq \text{ops}(\text{EST}_n, t_t)$.*

Proof. From Lemma ^{lem::establishedOrder}9 we know that $\text{ops}(\text{EST}_n, t_i) \subseteq \text{ops}(\text{EST}_n, t_t)$. So it remains to show the operations of $\text{ops}(\text{EST}_{n.\text{left}}, t_i) \cup \text{ops}(\text{EST}_{n.\text{right}}, t_i) - \text{ops}(\text{EST}_n, t_i)$, which we call new operations, are all in $\text{ops}(\text{EST}_n, t_t)$. If `TryAppend` returns true a block is appended to `n`. From the code of the `CreateBlock()` this block includes the established blocks in `n`'s children at t_i . Since the head in `CreateBlock()` is read after t_i . So the new operations are in the block appended to `n`. \square

Mention this new block is established in t_t . last sentence is not complete

Lemma 11 (^{Double Refresh}Double Refresh). *Consider two consecutive failed instances R_1, R_2 of `n.Refresh()` by some process. Let t_1 be the time R_1 is invoked and t_2 be the time R_2 terminated. We have $\text{ops}(\text{EST}_{n.\text{left}}, t_1) \cup \text{ops}(\text{EST}_{n.\text{right}}, t_1) \subseteq \text{ops}(\text{EST}_n, t_2)$.*

Proof.

If Line ^{cas}313 of R_1 or R_2 returns true, then the claim is held by Lemma ^{lem::trueRefresh}10. If not, then there is another successful instance R'_2 of `n.Refresh()` that has done `TryAppend()` successfully into `n.blocks[i+1]`. If R_2 reads some value greater than $i+1$ in Line ^{readSize}310 it means a successful instance of `Refresh()` started after Line ^{readSize}310 of R_1 and finished its Line ^{incrementHead1}318 or ^{incrementHead2}321 before ^{readSize}310 of R_2 , from Lemma ^{lem::trueRefresh}10 by the end of this instance $\text{ops}(\text{EST}_{n.\text{left}}, t_1) \cup \text{ops}(\text{EST}_{n.\text{right}}, t_1)$ has been propagated. In Figure 1 we see why the block R'_2 is appending contains established block in the `n`'s children at t_1 , since it creates a block reading the head after t_1 . \square

last sentence need more detail and should be earlier. define i and tell why R_2 exists

this chain with more depth should be in the proof

Corollary 12 (^{Propagate Step}Propagate Step). *All operations in `n`'s children's established blocks before line ^{firstRefresh}302 are guaranteed to be in `n`'s established blocks after line ^{secondRefresh}303.*

Proof. Lines ^{firstRefresh}302 and ^{secondRefresh}303 satisfy the preconditions of Lemma ^{doubleRefresh}11. \square

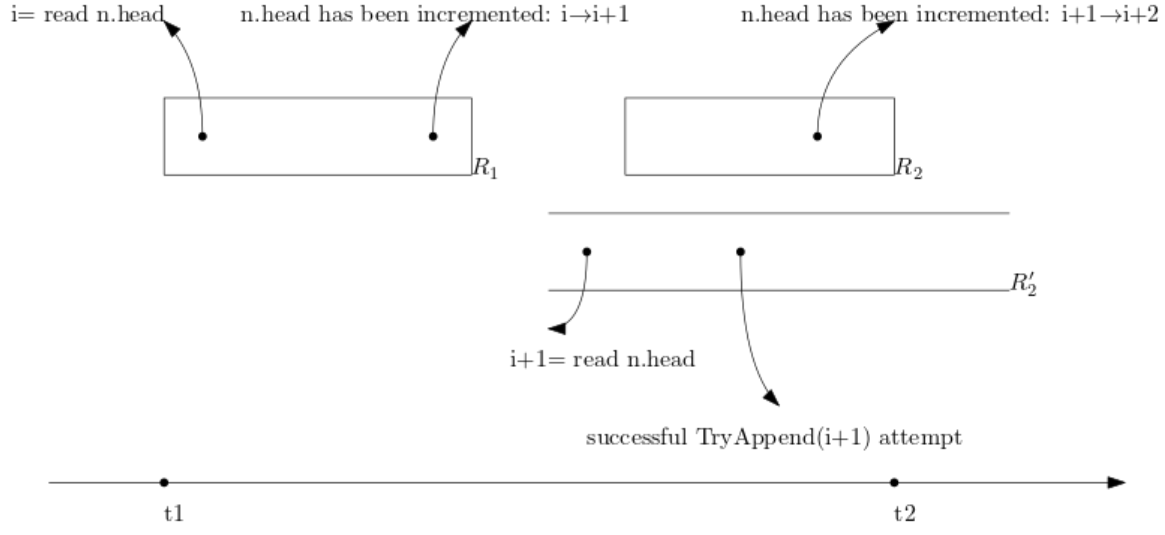


Figure 1: $t1 < r_1$ reading head $<$ incrementing $\mathbf{n.head}$ from i to $i+1 < R'_2$ reading head $<$ $\text{TryAppend}(i+1) <$ incrementing $\mathbf{n.head}$ from $i+1$ to $i+2 < t2$

Corollary 13. After $\text{Append}(\text{blk})$ finishes $\text{ops}(\text{blk}) \subseteq \text{ops}(\text{root.blocks}[\mathbf{x}])$ for some \mathbf{x} and only one \mathbf{x} .

Proof. Follows from Lemma 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000.

□

blockSize

Lemma 14 (Block Size Upper Bound). Each block contains at most one operation from each process.

Proof. By proof of contradiction, assume there are more than one operation from process p in block b in node n . A process cannot invoke more than one operations concurrently. From p 's operations in b , let op_1 be the first operation invoked and op_2 be the second one. Note that it is terminated before op_2 started. So before appending op_2 to the tree op_1 exists in every node from the path of p 's leaf to the root. So there is some block b' before b in n containing op_1 . op_1 existing in b and b' contradicts with 17. □

blocksBound

Lemma 15 (Subblocks Upperbound). Each block has at most p direct subblocks.

Proof. It follows directly from Lemma 14 and the observation that each block contains at least one operation, induced from Line 17. □

ordering

Definition 16 (Ordering of operations inside the nodes). ► Note that processes are numbered from 1 to p , left to right in the leaves of the tree and from Lemma 14 we know there is at most one operation from each process in a given block.

- We call operations strictly before op in the sequence of operations S , prefix of the op .
- $E(n, b)$ is the sequence of enqueue operations $\in \text{ops}(\mathbf{n.blocks}[b])$ ordered by their process id.
- $E(n, b, i)$ is the i th enqueue in $E(n, b)$.
- $D(n, b)$ is the sequence of dequeue operations $\in \text{ops}(\mathbf{n.blocks}[b])$ ordered by their process id.
- $D(n, b, i)$ is the i th dequeue in $D(n, b)$.
- Order of the enqueue operations in n : $E(n) = E(n, 1).E(n, 2).E(n, 3)...$
- $E(n, i)$ is the i th enqueue in $E(n)$.
- Order of the dequeue operations in n : $D(n) = D(n, 1).D(n, 2).D(n, 3)...$
- $D(n, i)$ is the i th dequeue in $D(n)$.
- Linearization: $L = E(\text{root}, 1).D(\text{root}, 1).E(\text{root}, 2).D(\text{root}, 2).E(\text{root}, 3).D(\text{root}, 3)...$

Note that in the non-root nodes we only order enqueues and dequeues among the operations of their own type. Since $\text{GetENQ}()$ only searches among enqueues and $\text{IndexDEQ}()$ works on dequeues.

get **Lemma 17** (Get correctness). *If $n.blocks[b].num_{enq} \geq i$ then $n.GetENQ(b, i)$ returns $E(n, b, i)$.*

Proof. We are going to prove this lemma by induction on the height of the tree. The base case for the leaves of the tree is pretty straight forward. Since leaf blocks contain exactly one operation then only $GetENQ(b, 1)$ can be called on leaves. $leaf.GetENQ(b, 1)$ returns the operation stored in the b th block of leaf l . For non leaf nodes in Line 404 it is decided that the i th enqueue in block b of internal node n resides in the left child or the right child of n . By definition of $E(n, b)$ operations from the left child come before the operations of the right child. Having sum_{enq} , the prefix sum of the number of enqueues we can compute the direct subblock containing the enqueue we are finding for with binary search. Then $n.child.GetENQ(\text{block containing}, \text{order in the block})$ is invoked which returns the correct operation by the hypothesis of the induction. \square

I'm not sure it is going to be long and boring to talk about the parameters, since the reader can find out them.

dsearch **Lemma 18** (DSearch correctness). *If $root.blocks[end].num_{enq} \geq i$ and $E(root, i)$ is the response to some $Dequeue()$ in $root.blocks[end]$ then $DSearch(i, end)$ returns b such that $root.blocks[b]$ contains $E(root, b, i)$ in $\Theta(\log(root.blocks[b].length + root.blocks[end].length))$ steps.*

Proof. First we show $end - b \leq root.blocks[b].length + root.blocks[end].length$. We know each block size is greater than 0. So every block in $root.blocks[b..end]$ contains at least one $Enqueue()$ or one $Dequeue()$. There cannot be more than $root.blocks[b].length$ $Dequeue()$ s in $root.blocks[b+1..end-1]$, since the queue would become empty after b th block end before end and $E(n, i)$ could not be the response to to some DEQ in end . And since the length of the queue would become $root.blocks[end].length$ in the end so there cannot be more than $root.blocks[end].length$ $Dequeus$ in $root.blocks[b..end]$. Cause if there was more then the end's length would become more than $root.blocks[end].length$.

Now that we know there are at most $root.blocks[b].length + root.blocks[end].length$ distance between end and b then with doubling search in $\log root.blocks[b].length + root.blocks[end].length$ steps we reach a block c that the $c.sum_{enq}$ is less than i and the distance between c and end is not more than $2 \times root.blocks[b].length + root.blocks[end].length$. So the binary search takes $\Theta(\log root.blocks[b].length + root.blocks[end].length)$ steps. \square

superBlock **Lemma 19** (Computing SuperBlock). *After computing line ^{computeSuper}418 of $n.IndexDEQ(b, i)$, $n.parent.blocks[superblock]$ contains $D(n, b, i)$.*

Proof. 1. Value read for $super[b.group]$ in line 418 is not null.

► Values c_{dir} read in lines 23, $super$ are set before incrementing in lines 26, 27.

2. $super[]$ preserves order from child to parent; if in a child block b is before c then $b.group \leq c.group$ and $super[b.group] \leq super[c.group]$

► Follows from the order of lines 60, 48, 49.

3. $super[i+1] - super[i] \leq p$

► In a Refresh with successful CAS in line 46, $super$ and $counter$ are set for each child in lines 48, 49. Assume the current value of the counter in node n is $i+1$ and still $super[i+1]$ is not set. If an instance of successful $Refresh(n)$ finishes $super[i+1]$ is set a new value and a block is added after $n.parent[super[i]]$. There could be at most p successful unfinished concurrent instances of $Refresh()$ that have not reached line 49. So the distance between $super[i+1]$ and $super[i]$ is less than p .

4. Superblock of b is within range $\pm 2p$ of the $super[b.time]$.

► $super[i]$ is the index of the superblock of a block containing block b , followed by Lemma ^{superCounter}23. It is trivial to see that $n.super$ and $n.b.counter$ are increasing. $super(b)$ is the real superblock of b . $super(t)$ is the index of the superblock of the last block with time t . If $b.time$ is t we have:

$$super[t] - p \leq super[t-1] \leq super(t-1) \leq super(b) \leq super(t+1) \leq super(t+1) \leq super[t] + p$$

□

Lemma 20 (Index correctness). $n.\text{IndexDEQ}(b, i)$ returns the rank in $D(\text{root})$ of $D(n, b, i)$.

Proof. We can see the base case $\text{root}.\text{IndexDEQ}(b, i)$ is trivial. $n.\text{IndexDEQ}(b, i)$ computes the superblock of the i th Dequeue in the b th block of n in $n.\text{parent}$ (Lemma 19). Then the order in $D(n.\text{parent}, \text{superblock})$ is computed and $\text{index}()$ is called on $n.\text{parent}$ recursively. It is easy to see why the second is correct. Correctness of computing superblock comes from Lemma 19. □

Do I need to talk about the computation of the order in the parent which is based on the definition of ordering of dequeues in a block?

search **Lemma 21** (Search Ranges). *Preconditions of all invocation of BSearch are satisfied.*

Proof. Line 83: $\text{Get}(i)$ is called if the result of a dequeue is not null. The search is among all blocks in the root.

Line 88: This search tries to find the i th enqueue, knowing that it is in the left child. Search is done over the left subblocks. The start and end of the range are followed by definition. Line 92 is the same.

Line 101: Here, the goal is to find the superblock. We know the distance between answer and the $\text{super}[i]$ is at most p , since at most p processes could die. □

computeHead **Lemma 22** (Computing Queue's Head block). *Let S be the state of an empty queue if the operations in prefix in L of i th dequeue in $D(\text{root}, b)$ are applied on it. $\text{FindResponse}()$ returns (b, i) which $E(\text{root}, b, i)$ is the the head of the queue in S . If the queue is empty in S then it returns $\langle -1, -- \rangle$.*

Proof. The length of the queue if the operations in the prefix for the b th block in the root are applied with the order of L is stored in the $\text{root.blocks}[b].\text{length}$. It is computed while creating the block in Line 349. If the length of a queue is greater than 0 then a $\text{Dequeue}()$ would decrease the length of the queue, otherwise the length of the queue remains 0. Having length of the queue after the previous block and number of enqueues and dequeues in the block, Line 349 computes whether the queue becomes empty or the length of it. □

HOW? How to prove mathematically that $\text{ax}(\text{root.blocks}[i-1].\text{length} + b.\text{num}_{\text{enq}} - b.\text{num}_{\text{deq}}, 0)$ is the length of the queue after the block. I can only explain it here.

erCounter **Lemma 23** (Validity of super and counter). *If $\text{super}[i] \neq \text{null}$, then $\text{super}[i]$ in node n is the index of the superblock of a block with $\text{time}=i \pm p$.*

rootRange **Lemma 24** (Root search range). *$\text{root.size} - \text{root.FindMostRecentDone}()$ is $O(p^2 + q)$, which p is # processes and q is the length of the queue.*

Theorem 25 (Main). *The queue implementation is linearizable.*

Lemma 26 (Time analysis). *$n.\text{GetEnq}(b, i)$, $n.\text{Index}(b, i)$ take $O(\log^2 p)$ steps. Search in the root may take $O(\log Q + p^2)$ steps. Helping is done every p^2 block appended to the root and takes $p \times \log^2 p$ steps. Amortized time consumed for helping by each process is $O(\log^2 p)$.*