IMP2 Binary Heap

Simon Griebel

October 13, 2025

Abstract

In this submission array-based binary minimum heaps are formalized. The correctness of the following heap operations is proven: insert, get-min, delete-min and make-heap. These are then used to verify an in-place heapsort. The formalization is based on IMP2, an imperative program verification framework implemented in Isabelle/HOL. The verified heap functions are iterative versions of the partly recursive functions found in “Algorithms and Data Structures – The Basic Toolbox” by K. Mehlhorn and P. Sanders and “Introduction to Algorithms” by T. H. Cormen, C. E. Leiserson, R. L. Rivest and C. Stein.

Contents

theory *IMP2-Binary-Heap*
  imports *IMP2.IMP2 IMP2.IMP2-Aux-Lemmas*
begin

1 Introduction

In this submission imperative versions of the following array-based binary minimum heap functions are implemented and verified: insert, get-min, delete-min, make-heap. The latter three are then used to prove the correctness of an in-place heapsort, which sorts an array in descending order. To do that in Isabelle/HOL, the proof framework IMP2 [2] is used. Here arrays are modeled by $int \Rightarrow int$ functions. The imperative implementations are iterative versions of the partly recursive algorithms described in [3] and [1].

This submission starts with the basic definitions and lemmas, which are needed for array-based binary heaps. These definitions and lemmas are parameterised with an arbitrary (transitive) comparison function (where such a function is needed), so they are not only applicable to minimum heaps. After some more general, useful lemmas on arrays, the imperative minimum heap functions and the heapsort are implemented and verified.

2 Heap Related Definitions and Theorems

2.1 Array Bounds

A small helper function is used to define valid array indices. Note that the lower index bound l is arbitrary and not fixed to 0 or 1. The upper index bound r is not a valid index itself, so that the empty array can be denoted by having $l = r$.

abbreviation *bounded* :: $int \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
  $bounded\ l\ r\ x \equiv l \leq x \wedge x < r$

2.2 Parent and Children

2.2.1 Definitions

For the notion of an array-based binary heap, the parent and child relations on the array indices need to be defined.

definition *parent* :: $int \Rightarrow int \Rightarrow int$ **where**
  $parent\ l\ c = l + (c - l - 1)\ div\ 2$

definition *l-child* :: $int \Rightarrow int \Rightarrow int$ **where**
  $l\text{-}child\ l\ p = 2 * p - l + 1$

definition *r-child* :: *int* ⇒ *int* ⇒ *int* **where**
r-child l p = *2* ∗ *p* − *l* + *2*

2.2.2 Lemmas

lemma *parent-upper-bound*: *parent l c* < *c* ⟷ *l* ≤ *c*
⟨*proof*⟩

lemma *parent-upper-bound-alt*: *l* ≤ *parent l c* ⟹ *parent l c* < *c*
⟨*proof*⟩

lemma *parent-lower-bound*: *l* ≤ *parent l c* ⟷ *l* < *c*
⟨*proof*⟩

lemma *grand-parent-upper-bound*: *parent l* (*parent l c*) < *c* ⟷ *l* ≤ *c*
⟨*proof*⟩

corollary *parent-bounds*: *l* < *x* ⟹ *x* < *r* ⟹ *bounded l r* (*parent l x*)
⟨*proof*⟩

lemma *l-child-lower-bound*: *p* < *l-child l p* ⟷ *l* ≤ *p*
⟨*proof*⟩

corollary *l-child-lower-bound-alt*: *l* ≤ *x* ⟹ *x* ≤ *p* ⟹ *x* < *l-child l p*
⟨*proof*⟩

lemma *parent-l-child*[*simp*]: *parent l* (*l-child l n*) = *n*
⟨*proof*⟩

lemma *r-child-lower-bound*: *l* ≤ *p* ⟹ *p* < *r-child l p*
⟨*proof*⟩

corollary *r-child-lower-bound-alt*: *l* ≤ *x* ⟹ *x* ≤ *p* ⟹ *x* < *r-child l p*
⟨*proof*⟩

lemma *parent-r-child*[*simp*]: *parent l* (*r-child l n*) = *n*
⟨*proof*⟩

lemma *smaller-l-child*: *l-child l x* < *r-child l x*
⟨*proof*⟩

lemma *parent-two-children*:
$(c = l\text{-}child\ l\ p \lor c = r\text{-}child\ l\ p) \longleftrightarrow parent\ l\ c = p$
$\langle proof \rangle$

2.3 Heap Invariants

2.3.1 Definitions

The following heap invariants and the following lemmas are parameterised with an arbitrary (transitive) comparison function. For the concrete function implementations at the end of this submission $\leq$ on ints is used.

For the *make-heap* function, which transforms an unordered array into a valid heap, the notion of a partial heap is needed. Here the heap invariant only holds for array indices between a certain valid array index m and r. The standard heap invariant is then simply the special case where $m = l$.

definition *is-partial-heap*
$:: ('a::order \Rightarrow 'a::order \Rightarrow bool) \Rightarrow (int \Rightarrow 'a::order) \Rightarrow int \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
$is\text{-}partial\text{-}heap\ cmp\ heap\ l\ m\ r = (\forall\ x.\ bounded\ m\ r\ x \longrightarrow$
$bounded\ m\ r\ (parent\ l\ x) \longrightarrow cmp\ (heap\ (parent\ l\ x))\ (heap\ x))$

abbreviation *is-heap*
$:: ('a::order \Rightarrow 'a::order \Rightarrow bool) \Rightarrow (int \Rightarrow 'a::order) \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
$is\text{-}heap\ cmp\ heap\ l\ r \equiv is\text{-}partial\text{-}heap\ cmp\ heap\ l\ l\ r$

During all of the modifying heap functions the heap invariant is temporarily violated at a single index i and it is then gradually restored by either *sift-down* or *sift-up*. The following definitions formalize these weakened invariants.
The second part of the conjunction in the following definitions states, that the comparison between the parent of i and each of the children of i evaluates to *True* without explicitly using the child relations.

definition *is-partial-heap-except-down*
$:: ('a::order \Rightarrow 'a::order \Rightarrow bool) \Rightarrow (int \Rightarrow 'a::order) \Rightarrow int \Rightarrow int \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
$is\text{-}partial\text{-}heap\text{-}except\text{-}down\ cmp\ heap\ l\ m\ r\ i = (\forall\ x.\ bounded\ m\ r\ x \longrightarrow$
$((parent\ l\ x \neq i \longrightarrow bounded\ m\ r\ (parent\ l\ x) \longrightarrow cmp\ (heap\ (parent\ l\ x))\ (heap\ x)) \land$
$(parent\ l\ x = i \longrightarrow bounded\ m\ r\ (parent\ l\ (parent\ l\ x))$
$\longrightarrow cmp\ (heap\ (parent\ l\ (parent\ l\ x)))\ (heap\ x))))$

abbreviation *is-heap-except-down*
:: $('a::order \Rightarrow 'a::order \Rightarrow bool) \Rightarrow (int \Rightarrow 'a::order) \Rightarrow int \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
is-heap-except-down cmp heap l r i $\equiv$ *is-partial-heap-except-down cmp heap l l r i*

As mentioned the notion of a partial heap is only needed for *make-heap*, which only uses *sift-down* internally, so there doesn't need to be an additional definition for the partial heap version of the *sift-up* invariant.

definition *is-heap-except-up*
:: $('a::order \Rightarrow 'a::order \Rightarrow bool) \Rightarrow (int \Rightarrow 'a::order) \Rightarrow int \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
is-heap-except-up cmp heap l r i $= (\forall\ x.\ bounded\ l\ r\ x \longrightarrow$
$((x \neq i \longrightarrow bounded\ l\ r\ (parent\ l\ x) \longrightarrow cmp\ (heap\ (parent\ l\ x))\ (heap\ x)) \wedge$
$(parent\ l\ x = i \longrightarrow bounded\ l\ r\ (parent\ l\ (parent\ l\ x))$
$\longrightarrow cmp\ (heap\ (parent\ l\ (parent\ l\ x)))\ (heap\ x))))$

2.3.2 Lemmas

lemma *empty-partial-heap*[*simp*]: *is-partial-heap cmp heap l r r*
⟨*proof*⟩

lemma *is-partial-heap-smaller-back*:
is-partial-heap cmp heap l m r $\Longrightarrow r' \leq r \Longrightarrow$ *is-partial-heap cmp heap l m r'*
⟨*proof*⟩

lemma *is-partial-heap-smaller-front*:
is-partial-heap cmp heap l m r $\Longrightarrow m \leq m' \Longrightarrow$ *is-partial-heap cmp heap l m' r*
⟨*proof*⟩

The second half of each array is a is a partial binary heap, since it contains only leafs, which are all trivial binary heaps.

lemma *snd-half-is-partial-heap*:
$(l + r)\ div\ 2 \leq m \Longrightarrow$ *is-partial-heap cmp heap l m r*
⟨*proof*⟩

lemma *modify-outside-partial-heap*:
assumes
$heap = heap'\ on\ \{m..<r\}$
is-partial-heap cmp heap l m r
shows *is-partial-heap cmp heap' l m r*

$\langle proof \rangle$

The next few lemmas formalize how the heap invariant is weakened, when the heap is modified in a certain way.

This lemma is used by *make-heap*.

lemma *partial-heap-added-first-el*:
assumes
$l \leq m$ $m \leq r$
is-partial-heap cmp heap l $(m + 1)$ r
shows *is-partial-heap-except-down cmp heap* l m r m
$\langle proof \rangle$

This lemma is used by *del-min*.

lemma *heap-changed-first-el*:
assumes *is-heap cmp heap* l r $l \leq r$
shows *is-heap-except-down cmp* $(heap(l := b))$ l r l
$\langle proof \rangle$

This lemma is used by *insert*.

lemma *heap-appended-el*:
assumes
is-heap cmp heap l r
$heap = heap'$ *on* $\{l..<r\}$
shows *is-heap-except-up cmp heap'* l $(r+1)$ r
$\langle proof \rangle$

2.3.3 First Heap Element

The next step is to show that the first element of the heap is always the "smallest" according to the given comparison function. For the proof a rule for strong induction on lower bounded integers is needed. Its proof is based on the proof of strong induction on natural numbers found in [4].

lemma *strong-int-gr-induct-helper*:
assumes $k < (i::int)$ $(\bigwedge i.\ k < i \Longrightarrow (\bigwedge j.\ k < j \Longrightarrow j < i \Longrightarrow P\ j) \Longrightarrow P\ i)$
shows $\bigwedge j.\ k < j \Longrightarrow j < i \Longrightarrow P\ j$
$\langle proof \rangle$

theorem *strong-int-gr-induct*:
assumes
$k < (i::int)$
$(\bigwedge i.\ k < i \Longrightarrow (\bigwedge j.\ k < j \Longrightarrow j < i \Longrightarrow P\ j) \Longrightarrow P\ i)$
shows $P\ i$

$\langle proof \rangle$

Now the main theorem, that the first heap element is the "smallest" according to the given comparison function, can be proven.

theorem *heap-first-el*:
  assumes
    $is\text{-}heap\ cmp\ heap\ l\ r$
    $transp\ cmp$
    $l < x\ x < r$
  shows $cmp\ (heap\ l)\ (heap\ x)$
  $\langle proof \rangle$

3 General Lemmas on Arrays

Some additional lemmas on *mset-ran*, *swap* and *eq-on* are needed for the final proofs.

3.1 Lemmas on *mset-ran*

abbreviation $arr\text{-}mset :: (int \Rightarrow 'a) \Rightarrow int \Rightarrow int \Rightarrow 'a\ multiset$ **where**
  $arr\text{-}mset\ arr\ l\ r \equiv mset\text{-}ran\ arr\ \{l..<r\}$

lemma *in-mset-imp-in-array*:
  $x \in\#\ (arr\text{-}mset\ arr\ l\ r) \longleftrightarrow (\exists\, i.\ bounded\ l\ r\ i \wedge arr\ i = x)$
  $\langle proof \rangle$

lemma *arr-mset-remove-last*:
  $l \leq r \Longrightarrow arr\text{-}mset\ arr\ l\ r = arr\text{-}mset\ arr\ l\ (r + 1) - \{\#arr\ r\#\}$
  $\langle proof \rangle$

lemma *arr-mset-append*:
  $l \leq r \Longrightarrow arr\text{-}mset\ arr\ l\ (r + 1) = arr\text{-}mset\ arr\ l\ r + \{\#arr\ r\#\}$
  $\langle proof \rangle$

corollary *arr-mset-append-alt*:
  $l \leq r \Longrightarrow arr\text{-}mset\ (arr(r := b))\ l\ (r + 1) = arr\text{-}mset\ arr\ l\ r + \{\#b\#\}$
  $\langle proof \rangle$

lemma *arr-mset-remove-first*:
  $i \leq r \Longrightarrow arr\text{-}mset\ arr\ (i - 1)\ r = arr\text{-}mset\ arr\ i\ r + \{\#arr\ (i - 1)\#\}$
  $\langle proof \rangle$

lemma *arr-mset-split*:

assumes $l \leq m$ $m \leq r$
shows *arr-mset arr* l r = *arr-mset arr* l m + *arr-mset arr* m r
⟨*proof*⟩

That the first element in a heap is the "smallest", can now be expressed using multisets.

corollary *heap-first-el-alt*:
assumes
transp cmp
is-heap cmp heap l r
$x \in\#$ (*arr-mset heap* l r)
heap $l \neq x$
shows *cmp* (*heap* l) x
⟨*proof*⟩

3.2 Lemmas on *swap* and *eq-on*

lemma *eq-on-subset*:
$arr1 = arr2$ *on* $R \Longrightarrow S \subseteq R \Longrightarrow arr1 = arr2$ *on* S
⟨*proof*⟩

lemma *swap-swaps*:
$arr' = swap\ arr\ x\ y \Longrightarrow arr'\ y = arr\ x \wedge arr'\ x = arr\ y$
⟨*proof*⟩

lemma *swap-only-swaps*:
$arr' = swap\ arr\ x\ y \Longrightarrow z \neq x \Longrightarrow z \neq y \Longrightarrow arr'\ z = arr\ z$
⟨*proof*⟩

lemma *swap-commute*: $swap\ arr\ x\ y = swap\ arr\ y\ x$
⟨*proof*⟩

lemma *swap-eq-on*:
$arr1 = arr2$ *on* $S \Longrightarrow x \notin S \Longrightarrow y \notin S \Longrightarrow arr1 = swap\ arr2\ x\ y$ *on* S
⟨*proof*⟩

corollary *swap-parent-eq-on*:
assumes
$arr1 = arr2$ *on* $- \{l..<r\}$
$l < c$ $c < r$
shows $arr1 = swap\ arr2$ (*parent* l c) c *on* $- \{l..<r\}$
⟨*proof*⟩

corollary *swap-child-eq-on*:

assumes
$arr1 = arr2\ on - \{l..<r\}$
$c = l\text{-}child\ l\ p \lor c = r\text{-}child\ l\ p$
$l \le p\ c < r$
shows $arr1 = swap\ arr2\ p\ c\ on - \{l..<r\}$
$\langle proof \rangle$

lemma *swap-child-mset*:
assumes
$arr\text{-}mset\ arr1\ l\ r = arr\text{-}mset\ arr2\ l\ r$
$c = l\text{-}child\ l\ p \lor c = r\text{-}child\ l\ p$
$l \le p\ c < r$
shows $arr\text{-}mset\ arr1\ l\ r = arr\text{-}mset\ (swap\ arr2\ p\ c)\ l\ r$
$\langle proof \rangle$

The following lemma shows, which propositions have to hold on the pre-swap array, so that a comparison between two elements holds on the post-swap array. This is useful for the proofs of the loop invariants of *sift-up* and *sift-down*. The lemma is kept quite general (except for the argument order) and could probably be more closely related to the parent relation for more concise proofs.

lemma *cmp-swapI*:
fixes $arr::'a::order \Rightarrow 'a::order$
assumes
$m < n \land x < y$
$m < n \land x < y \Longrightarrow x = m \Longrightarrow y = n \Longrightarrow P\ (arr\ n)\ (arr\ m)$
$m < n \land x < y \Longrightarrow x \neq m \Longrightarrow x \neq n \Longrightarrow y \neq m \Longrightarrow y \neq n \Longrightarrow P\ (arr\ m)\ (arr\ n)$
$m < n \land x < y \Longrightarrow x = m \Longrightarrow y \neq m \Longrightarrow y \neq n \Longrightarrow P\ (arr\ y)\ (arr\ n)$
$m < n \land x < y \Longrightarrow x = n \Longrightarrow y \neq m \Longrightarrow y \neq n \Longrightarrow P\ (arr\ m)\ (arr\ y)$
$m < n \land x < y \Longrightarrow x \neq m \Longrightarrow x \neq n \Longrightarrow y = n \Longrightarrow P\ (arr\ m)\ (arr\ x)$
$m < n \land x < y \Longrightarrow x \neq m \Longrightarrow x \neq n \Longrightarrow y = m \Longrightarrow P\ (arr\ x)\ (arr\ n)$
shows $P\ (swap\ arr\ x\ y\ m)\ (swap\ arr\ x\ y\ n)$
$\langle proof \rangle$

4 Imperative Heap Implementation

The following imperative heap functions are based on [3] and [1]. All functions, that are recursive in these books, are iterative in the following implementations. The function definitions are done with IMP2 [2]. From now on the heaps only contain ints and only use $\le$ as comparison function. The auxiliary lemmas used from now on are heavily modeled after the proof goals, that are generated by the vcg tool (also part of IMP2).

4.1 Simple Functions

4.1.1 Parent, Children and Swap

In this section the parent and children relations are expressed as IMP2 procedures. Additionally a simple procedure, that swaps two array elements, is defined.

procedure-spec *prnt* (l, x) **returns** p
  assumes *True*
    ensures $p = \mathit{parent}\ l_0\ x_0$
  defines ‹$p = ((x - l - 1)\ /\ 2 + l)$›
  ⟨*proof*⟩

procedure-spec *left-child* (l, x) **returns** lc
  assumes *True*
    ensures $lc = \mathit{l\text{-}child}\ l_0\ x_0$
  defines ‹$lc = 2 * x - l + 1$›
  ⟨*proof*⟩

procedure-spec *right-child* (l, x) **returns** rc
  assumes *True*
    ensures $rc = \mathit{r\text{-}child}\ l_0\ x_0$
  defines ‹$rc = 2 * x - l + 2$›
  ⟨*proof*⟩

procedure-spec *swp* $(heap, x, y)$ **returns** $heap$
  assumes *True*
    ensures $heap = \mathit{swap}\ heap_0\ x_0\ y_0$
  defines ‹$tmp = heap[x];\ heap[x] = heap[y];\ heap[y] = tmp$›
  ⟨*proof*⟩

4.1.2 *get-min*

In this section *get-min* is defined, which simply returns the first element (the minimum) of the heap. For this definition an additional theorem is proven, which enables the use of *Min-mset* in the postcondition.

theorem *heap-minimum*:
  assumes
    $l < r$
    $\mathit{is\text{-}heap}\ (\leq)\ heap\ l\ r$

shows *heap* l = *Min-mset* (*arr-mset heap* l r)
⟨*proof*⟩

procedure-spec *get-min* (*heap*, l, r) **returns** *min*
assumes $l < r \wedge$ *is-heap* ($\leq$) *heap* l r
ensures *min* = *Min-mset* (*arr-mset* $heap_0$ l_0 r_0)
for *heap*[] l r
defines ‹*min* = *heap*[l]›
⟨*proof*⟩

4.2 Modifying Functions

4.2.1 *sift-up* and *insert*

The next heap function is *insert*, which internally uses *sift-up*. In the beginning of this section *sift-up-step* is proven, which states that each *sift-up* loop iteration correctly transforms the weakened heap invariant. For its proof two additional auxiliary lemmas are used. After *sift-up-step* *sift-up* and then *insert* are verified.

sift-up-step can be proven directly by the smt-solver without auxiliary lemmas, but they were introduced to show the proof details. The analogous proofs for *sift-down* were just solved with smt, since the proof structure should be very similar, even though the *sift-down* proof goals are slightly more complex.

lemma *sift-up-step-aux1*:
fixes *heap*::*int* ⇒ *int*
assumes
is-heap-except-up ($\leq$) *heap* l r x
parent l $x \geq l$
(*heap* x) $\leq$ (*heap* (*parent* l x))
bounded l r k
$k \neq$ (*parent* l x)
bounded l r (*parent* l k)
shows (*swap heap* (*parent* l x) x (*parent* l k)) $\leq$ (*swap heap* (*parent* l x) x k)
⟨*proof*⟩

lemma *sift-up-step-aux2*:
fixes *heap*::*int* ⇒ *int*
assumes
is-heap-except-up ($\leq$) *heap* l r x
parent l $x \geq l$
heap $x \leq$ (*heap* (*parent* l x))

$bounded\ l\ r\ k$
$parent\ l\ k = parent\ l\ x$
$bounded\ l\ r\ (parent\ l\ (parent\ l\ k))$
shows
$swap\ heap\ (parent\ l\ x)\ x\ (parent\ l\ (parent\ l\ k)) \leq swap\ heap\ (parent\ l\ x)\ x\ k$
$\langle proof \rangle$

lemma *sift-up-step*:
fixes $heap::int \Rightarrow int$
assumes
$is\text{-}heap\text{-}except\text{-}up\ (\leq)\ heap\ l\ r\ x$
$parent\ l\ x \geq l$
$(heap\ x) \leq (heap\ (parent\ l\ x))$
shows $is\text{-}heap\text{-}except\text{-}up\ (\leq)\ (swap\ heap\ (parent\ l\ x)\ x)\ l\ r\ (parent\ l\ x)$
$\langle proof \rangle$

sift-up restores the heap invariant, that is only violated at the current position, by iteratively swapping the current element with its parent until the beginning of the array is reached or the current element is bigger than its parent.

procedure-spec *sift-up* $(heap,\ l,\ r,\ x)$ **returns** *heap*
assumes $is\text{-}heap\text{-}except\text{-}up\ (\leq)\ heap\ l\ r\ x \wedge bounded\ l\ r\ x$
ensures $is\text{-}heap\ (\leq)\ heap\ l_0\ r_0 \wedge$
$arr\text{-}mset\ heap_0\ l_0\ r_0 = arr\text{-}mset\ heap\ l_0\ r_0 \wedge$
$heap_0 = heap\ on - \{l_0..<r_0\}$
for $heap[]\ l\ x\ r$
defines ‹
$p = prnt(l,\ x);$
$while\ (x > l \wedge heap[x] \leq heap[p])$
$@variant$ ‹$x - l$›
$@invariant$ ‹$is\text{-}heap\text{-}except\text{-}up\ (\leq)\ heap\ l\ r\ x \wedge p = parent\ l\ x \wedge$
$bounded\ l\ r\ x \wedge arr\text{-}mset\ heap_0\ l_0\ r_0 = arr\text{-}mset\ heap\ l\ r \wedge$
$heap_0 = heap\ \ on - \{l..<r\}$›
{
$heap = swp(heap,\ p,\ x);$
$x = p;$
$p = prnt(l,\ x)$
}›
$\langle proof \rangle$

insert inserts an element into a heap by appending it to the heap and restoring the heap invariant with *sift-up*.

procedure-spec *insert* $(heap,\ l,\ r,\ el)$ **returns** $(heap,\ l,\ r)$

```
assumes is-heap (≤) heap l r ∧ l ≤ r
  ensures is-heap (≤) heap l r ∧
          arr-mset heap l r = arr-mset heap_0 l_0 r_0 + {#el_0#} ∧
          l = l_0 ∧ r = r_0 + 1  ∧ heap_0 = heap on − {l..<r}
for heap l r el
defines ‹
  heap[r] = el;
  x = r;
  r = r + 1;
  heap = sift-up(heap, l, r, x)
›
⟨proof⟩
```

4.2.2 *sift-down*, *del-min* and *make-heap*

The next heap functions are *del-min* and *make-heap*, which both use *sift-down* to restore/establish the heap invariant. *sift-down* is proven first (this time without additional auxiliary lemmas) followed by *del-min* and *make-heap*.

sift-down restores the heap invariant, that is only violated at the current position, by iteratively swapping the current element with its smallest child until the end of the array is reached or the current element is smaller than its children.

```
procedure-spec sift-down(heap, l, r, x) returns heap
  assumes is-partial-heap-except-down (≤) heap l x r x ∧ l ≤ x ∧ x ≤ r
    ensures is-partial-heap (≤) heap l_0 x_0 r_0 ∧
            arr-mset heap_0 l_0 r_0 = arr-mset heap l_0 r_0 ∧
            heap_0 = heap on − {l_0..<r_0}
  defines ‹
   lc = left-child(l, x);
   rc = right-child(l, x);
    while (lc < r ∧ (heap[lc] < heap[x] ∨ (rc < r ∧ heap[rc] < heap[x])))
      @variant ‹r − x›
      @invariant ‹is-partial-heap-except-down (≤) heap l x_0 r x ∧
                 x_0 ≤ x ∧ x ≤ r ∧ lc = l-child l x ∧ rc = r-child l x ∧
                 arr-mset heap_0 l r = arr-mset heap l r ∧
                 heap_0 = heap on − {l..<r}›
  {
    smallest = lc;
    if (rc < r ∧ heap[rc] < heap[lc]) {
      smallest = rc
    };
    heap = swp(heap, x, smallest);
    x = smallest;
```

$lc = left\text{-}child(l,\ x)$;
$rc = right\text{-}child(l,\ x)$
}›
⟨*proof*⟩

del-min needs an additional lemma which shows, that it actually removes (only) the minimum from the heap.

lemma *del-min-mset*:
fixes $heap::int \Rightarrow int$
assumes
$l < r$
$is\text{-}heap\ (\leq)\ heap\ l\ r$
$mod\text{-}heap = heap(l := heap\ (r - 1))$
$arr\text{-}mset\ mod\text{-}heap\ l\ (r - 1) = arr\text{-}mset\ new\text{-}heap\ l\ (r - 1)$
shows
$arr\text{-}mset\ new\text{-}heap\ l\ (r - 1) = arr\text{-}mset\ heap\ l\ r - \{\#Min\text{-}mset\ (arr\text{-}mset\ heap\ l\ r)\#\}$
⟨*proof*⟩

del-min removes the minimum element from the heap by replacing the first element with the last element, shrinking the array by one and subsequently restoring the heap invariant with *sift-down*.

procedure-spec *del-min* $(heap,\ l,\ r)$ **returns** $(heap,\ l,\ r)$
assumes $l < r \wedge is\text{-}heap\ (\leq)\ heap\ l\ r$
ensures $is\text{-}heap\ (\leq)\ heap\ l\ r \wedge$
$arr\text{-}mset\ heap\ l\ r = arr\text{-}mset\ heap_0\ l_0\ r_0 - \{\#Min\text{-}mset\ (arr\text{-}mset\ heap_0\ l_0\ r_0)\#\} \wedge$
$l = l_0 \wedge r = r_0 - 1 \wedge$
$heap_0 = heap\ on - \{l_0..<r_0\}$
for $heap\ l\ r$
defines ‹
$r = r - 1$;
$heap[l] = heap[r]$;
$heap = sift\text{-}down(heap,\ l,\ r,\ l)$
›
⟨*proof*⟩

make-heap transforms an arbitrary array into a heap by iterating through all array positions from the middle of the array up to the beginning of the array and calling *sift-down* for each one.

procedure-spec *make-heap* $(heap,\ l,\ r)$ **returns** *heap*
assumes $l \leq r$
ensures $is\text{-}heap\ (\leq)\ heap\ l_0\ r_0 \wedge$
$arr\text{-}mset\ heap\ l_0\ r_0 = arr\text{-}mset\ heap_0\ l_0\ r_0 \wedge$

$heap_0 = heap\ on - \{l_0..< r_0\}$
for *heap*[] *l r*
defines ‹
$y = (r + l)/2 - 1;$
$while\ (y \geq l)$
@*variant* ‹$y - l + 1$›
@*invariant* ‹$is\text{-}partial\text{-}heap\ (\leq)\ heap\ l\ (y + 1)\ r \wedge$
$arr\text{-}mset\ heap\ l\ r = arr\text{-}mset\ heap_0\ l_0\ r_0 \wedge$
$l - 1 \leq y \wedge y < r \wedge heap_0 = heap\ on - \{l..<r\}$›
{
$heap = sift\text{-}down(heap,\ l,\ r,\ y);$
$y = y - 1$
}›
⟨*proof*⟩

4.3 Heapsort Implementation

The final part of this submission is the implementation of the in-place heapsort. Firstly it builds the ≤-heap and then it iteratively removes the minimum of the heap, which is put at the now vacant end of the shrinking heap. This is done until the heap is empty, which leaves the array sorted in descending order.

4.3.1 Auxiliary Lemmas

Firstly the notion of a sorted array is needed. This is more or less the same as *ran-sorted* generalized for arbitrary comparison functions.

definition $array\text{-}is\text{-}sorted :: (int \Rightarrow int \Rightarrow bool) \Rightarrow (int \Rightarrow int) \Rightarrow int \Rightarrow int \Rightarrow bool$ **where**
$array\text{-}is\text{-}sorted\ cmp\ a\ l\ r \equiv \forall\, i.\ \forall\, j.\ bounded\ l\ r\ i \longrightarrow bounded\ l\ r\ j \longrightarrow i < j \longrightarrow cmp\ (a\ i)\ (a\ j)$

This lemma states, that the heapsort doesn't change the elements contained in the array during the loop iterations.

lemma *heap-sort-mset-step*:
fixes $arr{::}int \Rightarrow int$
assumes
$l < m\ m \leq r$
$arr\text{-}mset\ arr'\ l\ (m - 1) = arr\text{-}mset\ arr\ l\ m - \{\#Min\text{-}mset\ (arr\text{-}mset\ arr\ l\ m)\#\}$
$arr = arr'\ on - \{l..<m\}$
$mod\text{-}arr = arr'(m - 1 := Min\text{-}mset\ (arr\text{-}mset\ arr\ l\ m))$
shows $arr\text{-}mset\ arr\ l\ r = arr\text{-}mset\ mod\text{-}arr\ l\ r$

$\langle proof \rangle$

This lemma states, that each loop iteration leaves the growing second half of the array sorted in descending order.

lemma *heap-sort-second-half-sorted-step*:
  fixes $arr::int \Rightarrow int$
  assumes
    $l_0 < m$ $m \leq r_0$
    $arr = arr'$ $on - \{l_0..<m\}$
    $\forall i. \forall j.$ $bounded\ m\ r_0\ i \longrightarrow bounded\ m\ r_0\ j \longrightarrow\ i < j \longrightarrow arr\ j \leq arr\ i$
    $\forall x \in \#arr\text{-}mset\ arr\ l_0\ m.\ \forall y \in \#arr\text{-}mset\ arr\ m\ r_0.\ \neg\ x < y$
    $bounded\ (m - 1)\ r_0\ i$
    $bounded\ (m - 1)\ r_0\ j$
    $i < j$
    $mod\text{-}arr = (arr'(m - 1 := Min\text{-}mset\ (arr\text{-}mset\ arr\ l_0\ m)))$
  shows $mod\text{-}arr\ j \leq mod\text{-}arr\ i$
$\langle proof \rangle$

The following lemma shows that all elements in the first part of the array (the binary heap) are bigger than the elements in the second part (the sorted part) after every iteration. This lemma and the invariant of the *heap-sort* loop use $\neg\ x < y$ instead of $x \geq y$ since *vcg-cs* doesn't terminate in the latter case.

lemma *heap-sort-fst-part-bigger-snd-part-step*:
  fixes $arr::int \Rightarrow int$
  assumes
    $l_0 < m$
    $m \leq r_0$
    $arr\text{-}mset\ arr'\ l_0\ (m - 1) = arr\text{-}mset\ arr\ l_0\ m - \{\#Min\text{-}mset\ (arr\text{-}mset\ arr\ l_0\ m)\#\}$
    $arr = arr'$ $on - \{l_0..<m\}$
    $\forall x \in \#arr\text{-}mset\ arr\ l_0\ m.\ \forall y \in \#arr\text{-}mset\ arr\ m\ r_0.\ \neg\ x < y$
    $mod\text{-}arr = arr'(m - 1 := Min\text{-}mset\ (arr\text{-}mset\ arr\ l_0\ m))$
    $x \in \#arr\text{-}mset\ mod\text{-}arr\ l_0\ (m - 1)$
    $y \in \#arr\text{-}mset\ mod\text{-}arr\ (m - 1)\ r_0$
  shows $\neg\ x < y$
$\langle proof \rangle$

4.3.2 Implementation

Now finally the correctness of the *heap-sort* is shown. As mentioned, it starts by transforming the array into a minimum heap using *make-heap*. Then in each iteration it removes the first element from the heap with *del-min* after its value was retrieved with *get-min*. This value is then put at the position freed by *del-min*.

```
program-spec heap-sort
  assumes l ≤ r
    ensures array-is-sorted (≥) arr l₀ r₀ ∧
             arr-mset arr₀ l₀ r₀ = arr-mset arr l₀ r₀ ∧
             arr₀ = arr on − {l₀ ..<r₀ } ∧ l = l₀ ∧ r = r₀
  for l r arr[]
  defines ‹
    arr = make-heap(arr, l, r);
    m = r;
    while (m > l)
      @variant ‹m − l + 1›
      @invariant ‹is-heap (≤) arr l m ∧
        array-is-sorted (≥) arr m r₀ ∧
        (∀ x ∈# arr-mset arr l₀ m. ∀ y ∈# arr-mset arr m r₀. ¬ x < y) ∧
        arr-mset arr₀ l₀ r₀ = arr-mset arr l₀ r₀ ∧
        l ≤ m ∧ m ≤ r₀ ∧ l = l₀ ∧ arr₀ = arr on − {l₀ ..<r₀}›
    {
      min = get-min(arr, l, m);
      (arr, l, m) = del-min(arr, l, m);
      arr[m] = min
    }
  ›
  ⟨proof⟩
end
```

References

[1] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. Introduction to Algorithms. 2009.

[2] P. Lammich and S. Wimmer. IMP2 – Simple Program Verification in Isabelle/HOL. *Archive of Formal Proofs*, Jan. 2019. http://isa-afp.org/entries/IMP2.html, Formal proof development.

[3] K. Mehlhorn and P. Sanders. Algorithms and Data Structures: The Basic Toolbox. 2007.

[4] T. Nipkow, L. C. Paulson, and M. Wenzel. *Isabelle/HOL — A Proof Assistant for Higher-Order Logic*, page 184. 2018. http://isabelle.in.tum.de/dist/Isabelle2018/doc/tutorial.pdf.