

1st Asian-Pacific Summer School on Formal Methods

Course 12: Deductive verification of C programs with Frama-C and Jessie

Virgile Prevosto

CEA List

August 30, 2009



- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



long ra
 for 0 ->
 C1) if m
 tmp2 =
 of the

tmp2[0] = 1 << (n0 - 1) else 1 tmp[0] >> 1 << (n0 - 1) tmp2[0] = 1 << (n0 - 1) else tmp2[0] = tmp[0]; /* Then the second part takes the first one: if (n0 - 1) tmp[0] = 0; k = 0; k++ tmp[0][k] += mc2[0][k] * tmp2[k]; /* The [i][j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1. i = 1; tmp[0][i] >> 1; */ Final rounding: tmp2[0][i] is now represented on 9 bits: *if (tmp[0][i] < -256) m2[0][i] = -256; else if (tmp[0][i] > 255) m2[0][i] = 255; else m2[0][i] = tmp[0][i];

- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



long ra
 for 0 ->
 C1) if m
 tmp2 =
 of the

tmp2[0] = 1 << (nbl - 1) else 1 tmp[0] >> 1 << (nbl - 1) tmp[0] = 1 << (nbl - 1) + tmp2[0] - tmp[0]; /* Then the second part takes the first one... */
 tmp[0] = 0; k = 0; k = k + 1; tmp[0] = mc2[0][k] * tmp2[k]; /* The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1
 i = 1; tmp[0] >> 1; */ Final rounding: tmp2[0] is now represented on 9 bits. *if (tmp[0] < -256) m2[0] = -256; else if (tmp[0] > 255) m2[0] = 255; else m2[0] = tmp[0];

- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)

long ra
for 0 <=
C1) if (m
tmp2 =
of the

tmp2[0] = 1 << (n-1) / 2; else if (tmp1[0] >= 1 << (n-1) / 2) tmp2[0] = tmp1[0]; /* Then the second part takes the first one. */
tmp1[0] = 0; k = 0; k = k + 1; tmp1[0] += m2[0][k] * tmp2[k]; /* The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1
i = 1; tmp1[0] >= 1 << 2; /* Final rounding: tmp2[0] is now represented on 9 bits. */ if (tmp1[0] < -256) m2[0] = -256; else if (tmp1[0] > 255) m2[0] = 255; else m2[0] = tmp1[0];



- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



```

long ra
for 0 <= k <= 5
  C1: if (m
  tmp2 =
  of the
  tmp2[0] = 1 << (n0 - 1) else if (tmp1[0]) >= 1 << (n0 - 1) else tmp2[0] = tmp1[0]; /* Then the second part takes the first one
  tmp1[0] = 0; k = k + 1; tmp1[0] = m2[0][k] * tmp2[0]; /* The [k] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1
  i = i + 1; tmp1[0] >= 1 << i; /* Final rounding: tmp2[0] is now represented on 3 bits. *if (tmp1[0] < -256) m2[0] = -256; else if (tmp1[0] > 255) m2[0] = 255; else m2[0] = tmp1[0];
  
```

- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



long n
for 0 <=
C1) if m
tmp2 =
of the

tmp2[0] = 1 << (n-1) else 0 tmp[0] >> 1 << (n-1) tmp[0] = tmp[0] + tmp2[0] / * Then the second part takes the first part
tmp[0] = 0; k = 0; while (tmp[0] != m-2) { tmp[0] = m-2 - tmp[0]; k++; } The [k] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1
= 1 - tmp[0][0] >> 1; * Final rounding: tmp2[0] is now represented on 3 bits: if (tmp[0] < -256) m2[0] = -256; else if (tmp[0] > 255) m2[0] = 255; else m2[0] = tmp[0];

- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



long ra
 t for 0
 ct) if m
 tmp2 =
 e of the

tmp2[0] = 1 << (n-1) * 8; else if (tmp1[0] >= 1) << (n-1) * 8; tmp2[0] = tmp1[0] * 7; Then the second part takes the first bit of
 tmp1[0] = 0; k = 8; k++) tmp1[0] = mc2[0][k] * tmp2[0]; /* The [i][j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1 *MC1
 i = 1; tmp1[0] >>= 1; */ Final rounding: tmp2[0] is now represented on 9 bits. *if (tmp1[0] < -256) tmp2[0] = -256; else if (tmp1[0] > 255) tmp2[0] = 255; else tmp2[0] = tmp1[0];

- ▶ Hoare-logic based plugin, developed at INRIA Saclay.
- ▶ Input: a program and a specification
- ▶ Jessie generates **verification conditions**
- ▶ Use of **Automated Theorem Provers** to discharge the VCs
- ▶ If all VCs are proved, **the program is correct** with respect to the specification
- ▶ Otherwise: need to investigate why the proof fails
 - ▶ Fix bug in the code
 - ▶ Adds additional annotations to help ATP
 - ▶ Interactive Proof (Coq)



Usage

- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care

```

(long ra
for (i = 0; i < n; i++)
    C[i] = 0;
tmp2 =
of the

```

```

tmp2[i] = (i < (n-1)) ? tmp1[i] : 0;
tmp1[i] = 0; k = i;
i = 1; tmp1[i] >>= 1;

```

Final rounding: tmp2[i] is now represented on 9 bits. If (tmp1[i] < -256) m2[i] = -256; else if (tmp1[i] > 255) m2[i] = 255; else m2[i] = tmp1[i].



Usage

- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care



```

(long ra
for (i = 0
C1) if (m
tmp2 =
of the

```

```

tmp2[j] = (i < (n1 - 1) ? (i + 1) : (i - 1)) * tmp1[j] + (i < (n1 - 1) ? (i + 1) : (i - 1)) * tmp1[j];
tmp1[j] = 0; k = k + 1; tmp1[j] += m2[j][k] * tmp2[k];
The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1*MC1
i = i + 1; tmp1[j] += 1;
Final rounding: tmp2[j] is now represented on 9 bits. *M (tmp1[j] < -256) ? -256 : else if (tmp1[j] > 255) m2[j] = 255; else m2[j] = tmp1[j];

```

Usage

- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care



long ra
for 0 =>
C1) if (a
tmp2 =
of the

tmp2[0] = (l << (nbl - 1)) else if (tmp1[0]) >> (l << (nbl - 1)) tmp2[0] = (l << (nbl - 1)) + (tmp1[0] <> tmp1[0]); /* Then the second part takes the first one
tmp1[0][k] = 0; k = 8; k++) tmp1[0][k] += mc2[0][k] * tmp2[0][k]; /* The [i][j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1 *MC1
l = 1; tmp1[0][l] >> 1; /* Final rounding: tmp2[0][l] is now represented on 9 bits. *if (tmp1[0][l] < -256) tmp2[0][l] = -256; else if (tmp1[0][l] > 255) tmp2[0][l] = 255; else tmp2[0][l] = tmp1[0][l];

Usage

- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care



Usage

- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care



long n;
 for (i = 0; i < n; i++)
 {
 tmp2 = i * i;
 // ...
 }
 // ...
 tmp2[i] = i * i + (n - i) * i;
 // ...
 tmp2[i] = 0; k = k + 1; tmp2[i] = m * 2[i] * tmp2[k];
 // ...
 // Final rounding: tmp2[i] is now represented on 9 bits. If (tmp2[i] < -256) tmp2[i] = -256; else if (tmp2[i] > 255) tmp2[i] = 255; else tmp2[i] = tmp2[i];

Usage

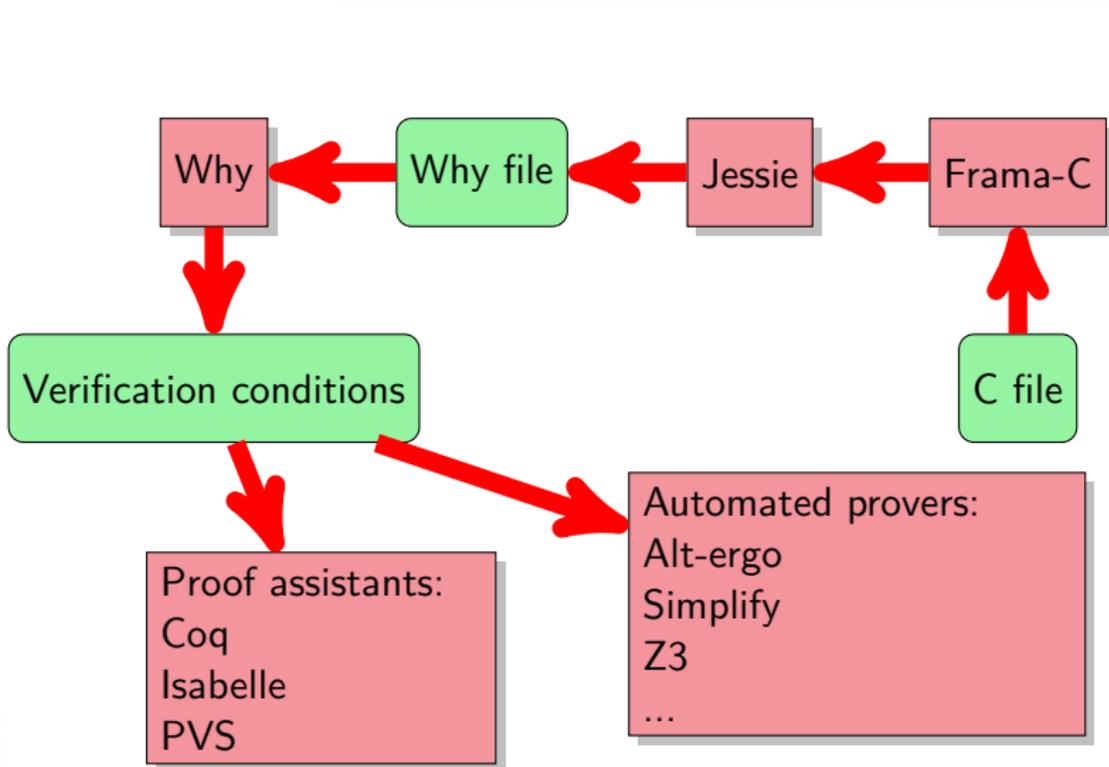
- ▶ Proof of functional properties of the program
- ▶ Modular verification (function per function)

Limitations

- ▶ Cast between pointers and integers
- ▶ Limited support for union type
- ▶ Aliasing requires some care



From Frama-C to Theorem Provers



Check safety of a function

- ▶ Pointer accesses
- ▶ Arithmetic overflow
- ▶ Division

```
unsigned int M;
```

```
void mean(unsigned int* p, unsigned int* q) {
    M = (*p + *q) / 2;
}
```



Jessie Usage

Function Contracts

Advanced Specification

Example 1: Searching

Example 2: Sorting

```

(long n)
for (i = 0; i < n; i++)
    C[i] = 0;
tmp2 = ...
// ...

```

```

tmp2[j][i] = (i < (n1 - 1)) ? tmp1[j][i] : ((i < (n1 - 1)) ? tmp2[j][i] : tmp1[j][i]); /* Then the second part looks like the first one:
tmp1[j][i] = 0; k = 0; k++ tmp1[j][k] += mc2[j][k] * tmp2[k][i]; /* The [j][i] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(M1 * M1) = MC2 * M1 * M1
i = 1; tmp1[j][i] >= 1; /* Final rounding: tmp2[j][i] is now represented on 9 bits. *if (tmp1[j][i] < -256) m2[j][i] = -256; else if (tmp1[j][i] > 255) m2[j][i] = 255; else m2[j][i] = tm

```



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of ACSL, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in

long ra
for 0 =>
C1) if (m
tmp2 =
of the

tmp2[j] = 0; k = (n-1) - j; else if (tmp1[j]) >= 0 { k = (n-1) - j; else tmp2[j] = tmp1[j]; } /* Then the second part copies the first part...
tmp1[0] = 0; k = 0; k++ tmp1[0][j] += m2[0][k] * tmp2[k][j]; /* The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(M1 * M1) = MC2*M1 * M1
i = 1; tmp1[0][i] >= 1; /* Final rounding: tmp2[0][i] is now represented on 9 bits: *if (tmp1[0][i] < -256) m2[0][i] = -256; else if (tmp1[0][i] > 255) m2[0][i] = 255; else m2[0][i] = tm



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of ACSL, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in

long ra
for 0 <=
C1) if (m
tmp2 =
of the

tmp2[j] = 0; for (k = 0; k <= 5; k++) tmp1[j][k] += m2[0][k] * tmp2[k][j]; /* The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(M1)*M1 = MC2*(M1)*M1 = 3 * tmp1[0][j] >>= 3. */ Final rounding: tmp2[0][j] is now represented on 9 bits: *if (tmp1[0][j] < -256) m2[0][j] = -256; else if (tmp1[0][j] > 255) m2[0][j] = 255; else m2[0][j] = tmp1[0][j];



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of ACSL, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in

long ra
for 0 <=
C1) if (m
tmp2 =
of the

tmp2[j] = 0; for (k = 0; k < n; k++) tmp1[j][k] += m2[j][k] * tmp2[k][j]; /* The [j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(MC1*M1) = MC2*M1 * MC1
i = 1; tmp1[j][i] >>= 1; */ Final rounding: tmp2[j][i] is now represented on 9 bits: if (tmp1[j][i] < -256) m2[j][i] = -256; else if (tmp1[j][i] > 255) m2[j][i] = 255; else tmp2[j][i] = m2[j][i];



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of ACSL, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of **ACSL**, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of **ACSL**, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of ACSL, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in



Safety of a program is important, but this is not sufficient: We also want it to do “the right thing”...

But in order for jessie to verify that, we need to explain it what “the right thing” is, and to explain it formally

This is the purpose of **ACSL**, ANSI/ISO C Specification Language.

- ▶ Behavioral specification language à la JML and Eiffel
- ▶ Function contracts
- ▶ Logic models
- ▶ Independent from any plug-in



- ▶ Functional specification
- ▶ Pre-conditions (requires)
- ▶ Post-conditions (ensures)

Example

```

unsigned int M;
/*@
  requires \valid(p) ^ \valid(q);
  ensures M ≡ (*p + *q) / 2;
*/
void mean(unsigned int* p, unsigned int* q) {
  if (*p ≥ *q) { M = (*p - *q) / 2 + *q; }
  else { M = (*q - *p) / 2 + *p; }
}

```



- ▶ Functional specification
- ▶ Pre-conditions (requires)
- ▶ Post-conditions (ensures)

Example

```

unsigned int M;
/*@
  requires \valid(p) ^ \valid(q);
  ensures M ≡ (*p + *q) / 2;
*/
void mean(unsigned int* p, unsigned int* q) {
  if (*p ≥ *q) { M = (*p - *q) / 2 + *q; }
  else { M = (*q - *p) / 2 + *p; }
}

```



- ▶ Functional specification
- ▶ Pre-conditions (requires)
- ▶ Post-conditions (ensures)

Example

```

unsigned int M;
/*@
  requires \valid(p) ^ \valid(q);
  ensures M ≡ (*p + *q) / 2;
*/
void mean(unsigned int* p, unsigned int* q) {
  if (*p ≥ *q) { M = (*p - *q) / 2 + *q; }
  else { M = (*q - *p) / 2 + *p; }
}

```



The specification:

```
/*@
```

```
  requires \valid(p) ^ \valid(q);
```

```
  ensures M ≡ (*p + *q) / 2;
```

```
  assigns M;
```

```
*/
```

```
void mean(unsigned int* p, unsigned int* q);
```



Jessie Usage

Function Contracts

Advanced Specification

Example 1: Searching

Example 2: Sorting

```

long n;
for (i = 0; i < n; i++)
    C[i] = 0;
tmp2 = ...
// ...

```

```

tmp2[0][i] = 0; // (i < (n-1) ? also if tmp1[i] >= 0 : (i < (n-1) ? tmp2[0][i] = (i <= (n-1) ? 1 : 0) : tmp2[0][i] = tmp1[0][i] ?) Then the second part looks like the first part.
tmp1[0][i] = 0; k = 5; k--> tmp1[0][i] += m2[0][k] * tmp2[0][i]; // The [i][j] coefficient of the matrix product MC2*TMP2, that is, *MC2*(TMP1) = MC2*(M1)*M1) = MC2*(M1)*M1
i = 1; tmp1[0][i] >= 1; // Final rounding: tmp2[0][i] is now represented on 9 bits. *if (tmp1[0][i] < -256) m2[0][i] = -256; else if (tmp1[0][i] > 255) m2[0][i] = 255; else m2[0][i] = tmp1[0][i];

```



Informal spec

- ▶ Input: a **sorted** array and its length, an element to search.
- ▶ Output: index of the element or -1 if not found

Implementation

```
int find_array(int* arr, int length, int query) {
    int low = 0;
    int high = length - 1;
    while (low ≤ high) {
        int mean = low + (high - low) / 2;
        if (arr[mean] ≡ query) return mean;
        if (arr[mean] < query) low = mean + 1;
        else high = mean - 1;
    }
    return -1;
}
```



Informal specification

- ▶ Input: an array and its length
- ▶ Output: the array is sorted in ascending order

```
int index_min(int* a, int low, int high);
```

```
void swap(int* arr, int i, int j);
```

```
void min_sort(int* arr, int length) {
    for(int i = 0; i < length; i++) {
        int min = index_min(arr,i,length);
        swap(arr,i,min);
    }
}
```

