

Vor und Nachname (Druckbuchstaben): _____

Legi Nummer: _____

252-0024-00L

Parallele Programmierung

ETH/CS: FS 2015

Basisprüfung

Montag, 10.08.2015

120 Minuten

Diese Prüfung enthält 22 Seiten (inklusive diesem Deckblatt) und 8 Aufgaben. Überprüfen Sie, dass keine Seiten fehlen. Füllen Sie alle oben verlangten Informationen aus. Schreiben Sie die Legi-Nummer oben auf jede einzelne Seite, für den Fall, dass Seiten verlorengehen oder abgetrennt werden.

Nehmen Sie sich am Anfang 5 Minuten Zeit, um alle Aufgaben durchzulesen. Während dieser Zeit ist es nicht erlaubt, Prüfungsfragen zu beantworten. Danach haben Sie 120 Minuten Zeit für die Lösung der Aufgaben.

Falls Sie sich durch irgendjemanden oder irgendetwas gestört fühlen, melden Sie dies sofort einer Aufsichtsperson. Wir sammeln die Prüfung zum Schluss ein. Wichtig: stellen Sie unbedingt selbst sicher, dass Ihre Prüfung von einem Assistenten eingezogen wird. Stecken Sie keine Prüfung ein und lassen Sie Ihre Prüfung nicht einfach am Platz liegen. Dasselbe gilt, wenn Sie früher abgeben wollen: bitte melden Sie sich lautlos und wir holen die Prüfung ab. Vorzeitige Abgaben sind nur bis 15 Minuten vor Prüfungsende möglich.

Wenn Sie zur Toilette müssen, melden Sie dies einer Aufsichtsperson durch Handzeichen. Es darf zur gleichen Zeit immer nur eine Studentin oder ein Student zur Toilette.

Es gelten die folgenden Regeln:

- **Lösungen müssen lesbar sein.** Verwenden Sie für Ihre Lösungen den verfügbaren Platz. Lösungen mit unklarer Reihenfolge oder anderweitig unverständlicher Präsentation können zu Punktabzügen führen.
- **Lösungen ohne Lösungsweg erhalten nicht die volle Punktzahl.** Eine korrekte Antwort ohne Lösungsweg, Erklärungen oder algebraischen Umformungen erhält keine Punkte; inkorrekte Antworten mit teilweise richtigen Formeln, Berechnungen und Umformungen können Teilpunkte erhalten.
- Falls mehr Platz benötigt wird, schreiben Sie auf die leeren Seiten der Prüfung oder fordern Sie bei den Assistenten zusätzliche Blätter an. Versehen Sie die Aufgabe mit einem klaren Hinweis, falls das der Fall ist. Als Richtlinie: **Die Aufgaben sollten sich alle in dem vorgegebenen Platz beantworten lassen.**
- Die Aufgaben können auf **Englisch oder Deutsch** beantwortet werden. Benutzen Sie keinen roten Stift!

Problem	Points	Score
1	4	
2	4	
3	18	
4	4	
5	26	
6	12	
7	12	
8	12	
Total:	92	

Java Sequential Programming (8 points)

1. Gegeben ist der folgende Java-Code.

Consider the following Java code.

```
interface Shape {
    public int area();
    public int perimeter();
}

class Square implements Shape {
    protected int length;
    public Square(int length) {
        this.length = length;
    }
    public int area() {
        return length * length;
    }
    public int perimeter () {
        return 4 * length;
    }
}

class Rectangle extends Square {
    int width;
    public Rectangle(int length, int width) {
        super(length);
        this.width = width;
    }
    public int area() {
        return length * width;
    }
}

public class Shapes {
    static void show(Shape x) {
        System.out.println(x.area() + ", " + x.perimeter());
    }

    public static void main(String[] args) {
        ...
    }
}
```

Für die Checkbox-Fragen gibt nur die korrekte Antwort 1 Punkt. Falsch markierte Checkboxes ergeben -1 Punkt. Die minimale Punktzahl einer Aufgabe ist 0.

For checkbox questions only the correct answer gives 1pt. Wrongly marked checkboxes result in -1 pts. The minimum number of points for one task is 0.

- (a) Sie fügen den folgenden Code in der `main`-Methode ein und führen dann das Programm aus. Was ist die erwartete Ausgabe? (1)
- ```
Shape a = new Rectangle(10, 20);
show(a);
```
- Das Programm kompiliert nicht.
  - Das Programm wirft eine Exception.
  - Das Programm gibt "200, 60" aus.
  - Das Programm gibt "200, 40" aus.
- You are inserting the following code in your main method and then run the program. What is the expected result?*
- The program does not compile.*
- The program throws an exception.*
- The program prints "200, 60".*
- The program prints "200, 40".*
- (b) Sie fügen den folgenden Code in der `main`-Methode ein und führen dann das Programm aus. Was ist die erwartete Ausgabe? (1)
- ```
Shape b = new Square(30);  
show(b);
```
- Das Programm kompiliert nicht.
 - Das Programm wirft eine Exception.
 - Das Programm gibt "900, 120" aus.
- You are inserting the following code in your main method and then run the program. What is the expected result?*
- The program does not compile.*
- The program throws an exception.*
- The program prints "900, 120".*
- (c) Sie fügen den folgenden Code in der `main`-Methode ein und führen dann das Programm aus. Was ist die erwartete Ausgabe? (1)
- ```
Square c = new Rectangle(10, 20);
show(c);
```
- Das Programm kompiliert nicht.
  - Das Programm wirft eine Exception.
  - Das Programm gibt "200, 60" aus.
  - Das Programm gibt "200, 40" aus.
  - Das Programm gibt "100, 40" aus.
- You are inserting the following code in your main method and then run the program. What is the expected result?*
- The program does not compile.*
- The program throws an exception.*
- The program prints "200, 60".*
- The program prints "200, 40".*
- The program prints "100, 40".*
- (d) Sie fügen den folgenden Code in der `main`-Methode ein und führen dann das Programm aus. Was ist die erwartete Ausgabe? (1)
- ```
Rectangle d = new Square(30);  
show(d);
```
- Das Programm kompiliert nicht.
 - Das Programm wirft eine Exception
 - Das Programm gibt "900, 120" aus.
- You are inserting the following code in your main method and then run the program. What is the expected result?*
- The program does not compile.*
- The program throws an exception.*
- The program prints "900, 120".*

2. Schreiben Sie eine Java-Funktion `sum` mit drei Argumenten: einem Integer-Array `a`, und zwei Integer-Werten `first` und `last`. Die Funktion soll die Summe aller Array-Werte mit gültigem Index von (inklusive) `first` bis (exklusive) `last` berechnen. Sie können `first ≤ last` annehmen. Die Summe ist 0 falls kein solches Element existiert. Die Funktion soll die berechnete Summe auf der Standard-Ausgabe ausgeben. Sie können die Methoden `min` und `max` verwenden, die jeweils zwei Argumente erhalten und das Minimum beziehungsweise Maximum der beiden Argumente zurück geben.

Write a Java function `sum` that takes three arguments: an array `a` of integers and two integers `first` and `last`. The function computes the sum of all array elements with valid indices between (and including) `first` and (excluding) `last`. You can assume `first ≤ last`. The sum is 0 if no such element exists. The function prints the sum to the standard output. You may assume the existence of the methods `min` and `max` that take as arguments two integers and return the minimum and respectively the maximum of their arguments. (4)

```
public static void sum(int[] a, int first, int last) {  
    .....  
    .....  
    .....  
    for (int i=.....; i<.....; .....) {  
        .....  
        .....  
        .....  
        .....  
    }  
    .....  
    .....  
}
```


Speedup, Amdahl, Gustafson (18 points)

3. Beantworten Sie die folgenden Fragen zum Thema Speedup sowie Amdahlsches und Gustafsonsches Gesetz:

(a) Schreiben Sie die Formel für den Speedup nach dem **Gustafsonschen** Gesetz auf und erläutern Sie kurz die Parameter.

Answer the following questions about speedup and both Amdahl's and Gustafson's laws:

*Write down the formula for speedup according to **Gustafson's** law and briefly describe its parameters.* (2)

.....
.....
.....
.....

(b) Die Auswertung eines Programms hat ergeben, dass ein Programm einen Speedup von 10 (skaliert) hat, wenn es auf 16 Prozessoren läuft.
Was ist der serielle Anteil nach dem **Gustafsonschen** Gesetz?

An evaluation of a program has shown a (scaled) speedup of 10 when running on 16 cores (2)

*What is the serial fraction according to **Gustafson's** law?*

.....
.....
.....
.....

(c) Was ist der serielle Anteil von (b) nach dem **Amdahlschen** Gesetz?

*What is the serial fraction of (b) according to **Amdahl's** law?* (2)

.....
.....
.....
.....

Synchronization (30 points)

4. Beschreiben Sie den Unterschied zwischen Parallelismus und Nebenläufigkeit so wie in der Vorlesung besprochen.

Describe the difference between parallelism and concurrency as it was discussed in the lectures. (4)

.....

.....

.....

.....

.....

5. Ein Zoo-Simulationsprogramm hat das folgende Problem: Es gibt zwei Thread-Klassen genannt Rabbit und Lion. Es gibt eine einzige Wasserquelle genannt WaterSource, welche folgendermassen verwendet werden muss:

In a Zoo simulation program, there exists the following problem: There are two classes of threads called Rabbit and Lion. There is a single water source called WaterSource that must be used in the following way:

- Tiere (Threads) unterschiedlicher Klasse dürfen nicht gleichzeitig von der Wasserquelle trinken.
- Jedes Tier (Thread), das von der Wasserquelle trinken will, kommt irgendwann zum trinken.

Animals (threads) of different type may not drink from the water source simultaneously. Every animal (thread) who needs to drink from the water source eventually succeeds.

Das Protokoll soll mit den Funktionen `enterLion()`, `leaveLion()`, `enterRabbit()`, `leaveRabbit()` implementiert werden. `enterLion()` blockiert den Caller bis es okay ist für einen Löwen zu trinken. `leaveLion()` wird aufgerufen wann immer ein Löwe fertig ist mit trinken. `enterRabbit()` und `leaveRabbit()` machen dasselbe für Rabbits. Als Beispiel,

The protocol is implemented via the following four functions `enterLion()`, `leaveLion()`, `enterRabbit()`, `leaveRabbit()`. `enterLion()` delays the caller until it is okay for a lion to drink. `leaveLion()` is called whenever a lion is finished drinking, while `enterRabbit()` and `leaveRabbit()` do the same for rabbits. For example,

```
waterSource.enterLion();
lion.drink(amount);
waterSource.leaveLion();
```

- (a) Nennen Sie zwei wichtige Konzepte, die ein korrektes paralleles Programm erfüllen sollte wenn Locks verwendet werden. *Name two important properties for the correct execution of parallel programs in the presence of locks?* (4)

.....

- (b) Implementieren Sie die Klasse WaterSource entsprechend der Spezifikation ohne dabei `synchronized (this){...}` zu benutzen. *Implement the class WaterSource according to the specification without using `synchronized (this){...}`.* (16)

```
public class WaterSource {
    // Lock to protect water source

    Lock waterLock = .....;

    // Monitor for notifications
    ..... monitor = .....;

    // Number of lions drinking at water source
    ..... int lionCount = 0;

    // Number of rabbits drinking at water source
    ..... int rabbitCount = 0;

    void enterLion() throws InterruptedException {
        .....;

        while (rabbitCount > 0) {
            .....;

            synchronized (monitor) {
                .....;
            }

            .....;
        }

        lionCount++;

        .....;
    }
}
```

```
void leaveLion() {  
  
    .....;  
  
    lionCount--;  
    if (lionCount == 0) {  
        synchronized (monitor) {  
            .....;  
        }  
    }  
  
    .....;  
}
```

```
void enterRabbit() throws InterruptedException {  
  
    .....;  
  
    while (lionCount > 0) {  
        .....;  
        synchronized (monitor) {  
            .....;  
        }  
        .....;  
    }  
    rabbitCount++;  
  
    .....;  
}
```

```
void leaveRabbit() {  
  
    .....;  
  
    rabbitCount--;  
    if (rabbitCount == 0) {
```

```
        synchronized (monitor) {  
            .....;  
        }  
    }  
    .....;  
}
```

(c) Erklären Sie, ob und weshalb Ihre Implementierung die beiden in Teilaufgabe a genannten Eigenschaften erfüllt.

Explain why your water source implementation does or does not satisfy the two properties given in part a of this question. (6)

.....
.....
.....
.....
.....
.....
.....
.....
.....

OpenMP & Data Parallelism (12 points)

6. Theoriefragen. Für die Checkbox-Fragen gibt nur die korrekte Antwort 1pt. Falsch markierte Checkboxen resultiert in -1pt. Die Minimale Punktzahl ist 0.

- (a) Was beschreibt das Programm im Data-Parallelism (was ist das Programmier-Modell)?
- Was getan werden muss (deklarativ).
 - Wie die Arbeit gemacht werden soll (imperativ).
 - Wie Daten-Objekte interagieren (objekt-orientiert).
- (b) Welche der folgenden Klassen von Patterns sind geeignet, um mit Data-Parallelism implementiert zu werden?
- Map: Führe eine Funktion auf jedem Element einer Liste/Sammlung aus
 - Komplexe Synchronisierung von Berechnungen
 - Reduktionen: Berechne die Summe, das Maximum/Minimum etc. einer Liste/Sammlung
 - Filter: Wähle einige Element einer Liste/Sammlung basierend auf einer Filterungs-Funktion, welche auf jedes Element angewendet wird, aus
 - Berechnungen mit komplexen Nachrichten-Patterns zwischen Tasks

Theory questions. For the checkbox questions only the correct answer gives 1pt. Wrongly marked checkboxes result in -1pts. The minimum number of points is 0.

What does the program describe in data parallelism (what is the programming model)? (1)

What needs to be done (declarative).

How the work should be done (imperative).

The interaction between data objects (object oriented).

Which of these patterns are well suited to be implemented using data parallelism? (3)

map: Execute a function on each element of a list/collection

Complex synchronization of computations

reductions: Calculate the sum/maximum/minimum/... of a list/collection.

filter: select some elements of a list/collection based on a filtering function which is applied to each element.

Computations with complex communication patterns between tasks

- (c) OpenMP Programmier-Frage: In dieser Frage sollen Sie einen parallelen Sortieralgorithmus implementieren, welcher wie folgt funktioniert: Zuerst sollen Sie mit dem Quicksort-Algorithmus n Teile eines Arrays parallel, und in-place sortieren (n ist die Anzahl Threads, welche von der OpenMP Laufzeitumgebung zur Verfügung gestellt wird), und dann alle sortierten Array-Teile mergen, um das sortierte Array zu erstellen. Sie können die Funktionen `merge` und `quicksort`, wie unten gezeigt, verwenden. Weiterhin können Sie annehmen, dass die OpenMP Laufzeitumgebung Ihnen eine Zweierpotenz von Threads zur Verfügung stellt und dass die Länge des Input-Arrays eine Zweierpotenz ist.

```
// from the OpenMP API:
// return the number of threads provided by the OpenMP
// runtime
int omp_get_num_threads();

// quicksort: sort elements [start, end) in the given array
// in-place.
void quicksort(int *array, int start, int end);

// merge: merge sorted elements [start, mid) and [mid, end)
// in the given array in-place
void merge(int *array, int start, int mid, int end);
```

Vervollständigen Sie das gegebene Programm auf der nächsten Seite.

OpenMP programming question: In this question, you should implement a parallel sort algorithm that works as follows: First use quicksort to sort n chunks of an array in parallel, and in-place (n being the number of threads provided by the OpenMP runtime) and then merge all the sorted chunks to produce the sorted array. You are given the functions `merge` and `quicksort` as shown below. You can assume that the OpenMP runtime will provide a power-of-two amount of threads and that the input array is power-of-two sized. (8)

Complete the program on the following page.

```
void parallel_sort(int *array, int length)
{
    // Number of parallel executions
    int runs = 0;
    // Number of elements to be sorted by each parallel thread
    int run_size = 0;

    #pragma omp parallel
    {

        runs = .....;

        run_size = .....;
        // Get index for current thread from 0 to
        // omp_get_num_threads()-1
        int thread_id = omp_get_thread_num();

        int start = .....;

        int end = .....;
        // sort interval [start, end)
        quicksort(array, start, end);
    }

    // recursively merge already sorted parts of the array
    // until whole array sorted
    while (runs > 1) {
        for (int i = 0; i < runs; i += 2) {

            merge(array, .....);
        }

        runs = .....;

        run_size = .....;
    }
}
```

OpenCL (12 points)

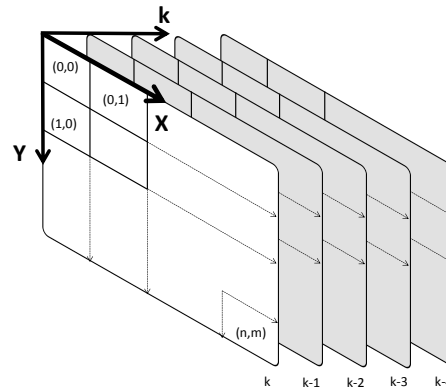


Abbildung 1

7. (a) Neben dem Studium arbeiten Sie an einem Bildbearbeitungsprogramm. Sie wollen einen Filter einbauen, der den Langzeitbelichtungseffekt (siehe Abb. 1) approximiert. Solch ein Effekt lässt sich erzielen, indem man 5 Bilder ohne Kamerabewegung aufnimmt und den Durchschnitt aus den 5 Bildern berechnet.

Implementieren Sie einen OpenCL-Kernel für den Zeit-Smoothing-Effekt. Der Kernel erhält als Input ein Schwarzweiss-Bild, repräsentiert durch einen Vektor von Integerwerten zwischen 0 und 255 (0 für Schwarz und 255 für Weiss). Der Kernel wird für jedes Pixel (x, y) des Output-Bildes aufgerufen und verarbeitet die 5 Eingabepixel (x, y, k) , $(x, y, k - 1)$, $(x, y, k - 2)$, $(x, y, k - 3)$ und $(x, y, k - 4)$.

- Die 5 Bilder werden als **ein** Vektor mit folgender Grösse gespeichert:

$$width \cdot height \cdot 5$$

- Den Index i für ein Pixel im Bildvektor mit den Koordinaten $[x, y, k]$ können sie mit der Formel berechnen:

$$i = x + y \cdot width + k \cdot (width \cdot height)$$

- Der gefilterte Pixelwert von Pixel (x, y) ist der Durchschnitt aus den 5 Eingabepixeln:

$$p(x, y) = \frac{1}{5} \sum_{i=1}^5 I_i(x, y)$$

Beside your studies you work on an image editing software. You want to approximate the long exposure effect with a filter in your program. You know that you can implement this effect by averaging 5 images taken from the exact same position. (8)

Implement the time-smoothing effect in an OpenCL-Kernel. The kernel takes a black and white image as input represented by a vector of integer values between 0 and 255 (0 for black and 255 for white). The kernel will be called for each pixel (x, y) of the output image and processes 5 input pixels (x, y, k) , $(x, y, k - 1)$, $(x, y, k - 2)$, $(x, y, k - 3)$ and $(x, y, k - 4)$.

*The 5 input images are stored as **one** vector with size:*

The index i for a pixel in the image vector with the coordinates $[x, y, k]$ can be computed with the formula:

The filtered pixel value at location (x, y) is the average of the 5 input-pixels:

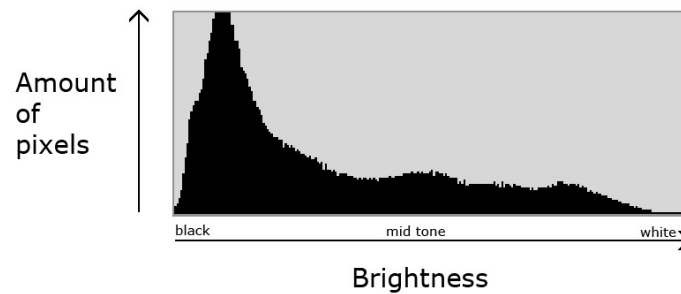


Abbildung 2

Der folgende Code stellt zwei alternative Implementierungen der Histogrammberechnung dar. N ist die Anzahl Pixel (Konstante).

The code below provides two alternative implementations of a histogram computation in Open CL. N is a constant holding the number of pixels.

Listing 1: Implementation A

```
kernel void hist(local int* A,
                 local int* B)
{
    int id = get_global_id(0);
    int t = A[id];
    atomic_inc(&B[t]);
}
```

Listing 2: Implementation B

```
kernel void hist(local int* A,
                 local int* B)
{
    int id = get_global_id(0);
    int t = A[id];
    for(int j=0; j<N; j++)
    {
        if(id == j)
            B[t]++;
        barrier();
    }
}
```

Analysieren Sie die beiden Kernel bezüglich ihrer Korrektheit (1pt) und bzgl. des Parallelismus (3pt) (und der daraus resultierenden Effizienz).

Read the code carefully and provide an analysis of the kernels in terms of correctness (1pt) and parallelism (3pt) (and the resulting efficiency).

.....

.....

.....

.....

.....

.....

.....

.....

Linearizability and Sequential Consistency(12 points)

8. In den folgenden Aufgaben seien A, B und C Threads, welche mit einem gemeinsam benutzten Stack-Objekt arbeiten. Die Spezifikation ist in der folgenden Interfacedefinition ersichtlich.

In the following questions let A, B and C denote threads operating on a shared stack object as specified in the listing below.

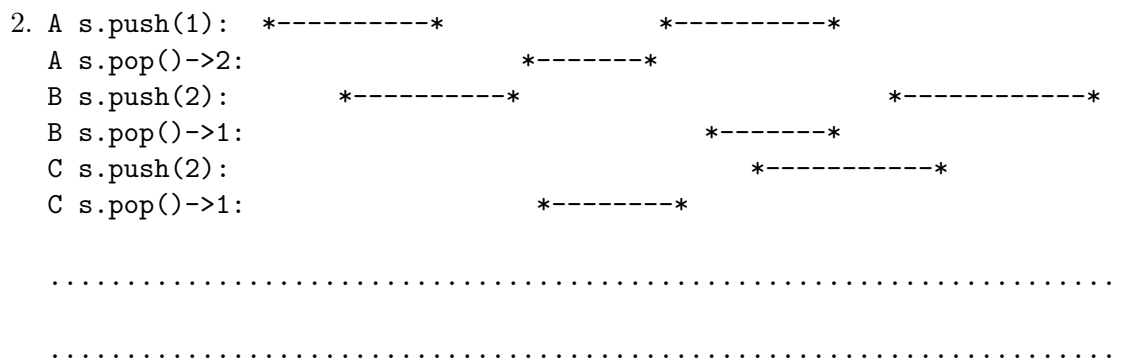
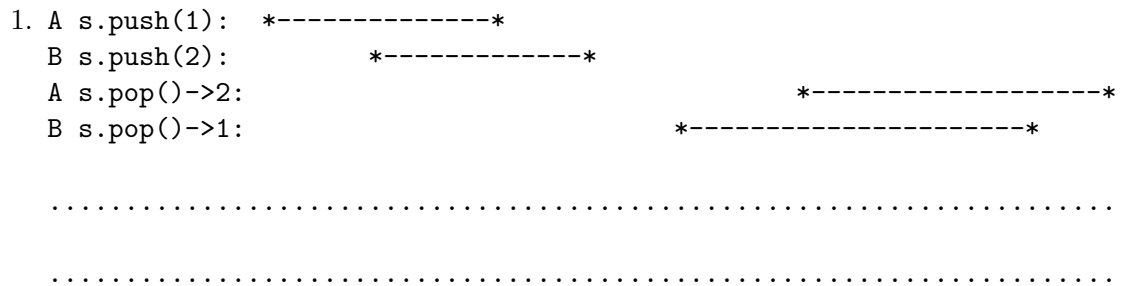
```
public interface Stack {
    /* pushes element v onto the stack */
    public void push(int v);

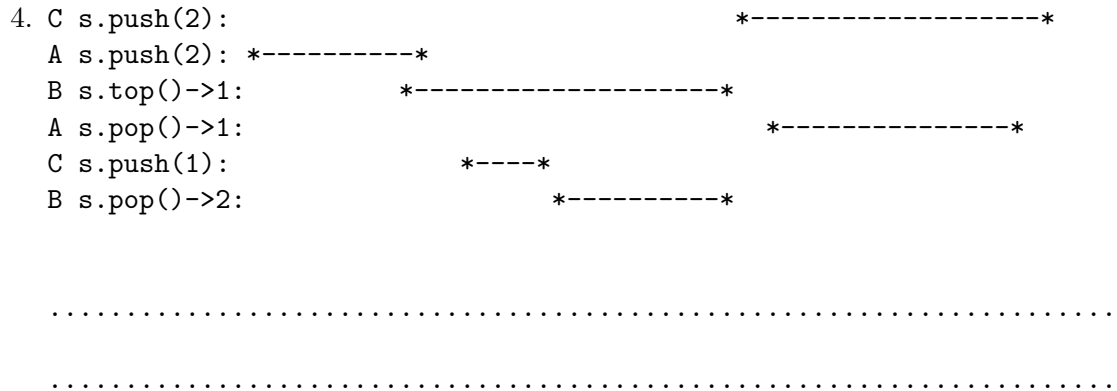
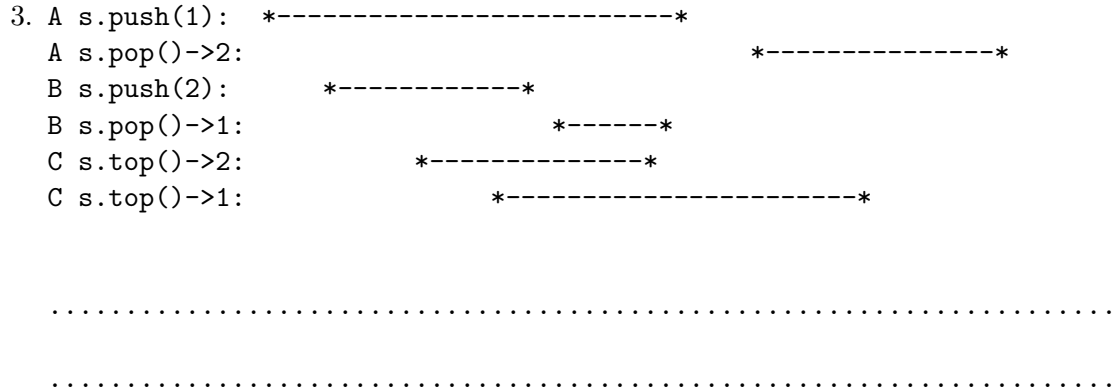
    /* removes the top element and returns it */
    public int pop();

    /* returns the top element */
    public int top();
}
```

(a) Welches der folgenden Szenarien ist linear konsistent? Markiere entweder den Linearisationspunkt oder erkläre, warum es nicht linear konsistent ist.

Which of the following scenarios are linearly consistent? Either mark the point of linearization or explain why it is not linearly consistent. (4)





(b) Gegeben sind folgende Historien H und G. Beantworten sie folgende Fragen für beide Historien. Begründen sie ihre Antworten. Richtig angekreuzte Antworten geben einen Punkt. Richtige Begründung liefert einen weiteren Punkt. Falsch angekreuzte Antworten geben einen Minuspunkt. Die minimale Punktzahl ist 0.

Consider the following two histories H and G. Answer the following questions for both histories. Justify your answer. Correct checkbox answers receive 1 point. Correct justification delivers another point. Wrong checkbox answers receive a negative point. The minimum number of points is 0. (6)

- | | |
|----------------|----------------|
| H= A s.push(x) | G= A s.push(x) |
| B s.pop() | A s:void |
| B s:x | A s.push(y) |
| A s:void | A s:void |
| A s.push(y) | B s.pop() |
| B s.push(y) | B s:x |
| A s:void | A s.pop |
| A s.pop | A s:y |
| A s:y | B s.push(y) |
| B s:void | B s:void |

Sind die Historien equivalent?

Yes No

Are these histories equivalent?

Yes No

(c)

.....
.....

Sind sie sequentiell?

H: Yes No

G: Yes No

Are they sequential?

H: Yes No

G: Yes No

(d)

.....
.....

Sind sie legal?

H: Yes No

G: Yes No

Are they legal?

H: Yes No

G: Yes No

(e)

.....
.....

(f) Ist die Komposition mehrerer sequentiell konsistenter Objekte selbst immer sequentiell konsistent? Belegen Sie Ihre Antwort mit einem (Gegen-)Beispiel.

Is the result of composing multiple sequentially consistent objects itself always sequentially consistent? Justify your answer with a (counter-)example. (2)

.....
.....
.....
.....
.....
.....
.....
.....
.....