

Gamaya Software developer assignment - v1.0

Gamaya Software Team
vlad.lapadatescu@gamaya.com
www.gamaya.com

November 30, 2016

1 Instructions

Thank you for taking the Gamaya Software Developer Assignment. The quizzes and assignments are meant to both evaluate some basic knowledge of several fields of interest for this position, but also to understand your particular strengths.

You should return your completed assignment within a week of reception. Please provide a final report in pdf format and package everything in a single file `SDEV3_NameSurname.zip`. If the package is larger than 10Mb please provide a link to download it.

2 Quiz

2.1 Java

1. You are given a class that can handle several operations on extremely large `byte` arrays. You are now required to update the class so that it can handle `int` and `float` arrays as well. Provide an example of how you would do it.
2. You are in the process of writing unit tests for class A. You want to test `someMethod` that receives an instance of class B as argument: `value == instanceA1.someMethod(instanceB2)`. It is extremely difficult to create an instance of `ClassB` in the test environment. What is the easiest method of writing this particular unit test for `ClassA`?
3. You have a fairly large (100+ classes) multithreaded Java application. You run it and it hangs. How do you go about debugging this application?
4. What is wrong with this code?

```
public static <T> T readTiffField(TiffDirectory directory, TagInfo tagInfo) {
    try {
        final TiffField field = findTiffField(directory, tagInfo);
        if (field == null) {
            return null;
        } else {
            if (!field.getValue().getClass().isArray())
                return (T) (field.getValue());
            else
                return null;
        }
    } catch (ImageReadException e) {
        LOGGER.error("There was an error trying to read field tag");
        return null;
    } catch (ClassCastException e) {
        LOGGER.error("Unable to properly cast the field in the required type");
        return null;
    }
}
```

¹Instance of class A

²Instance of class B

}

TiffField.getValue can either return a Short, Integer, Double or an array of those elements.

5. You produced a nice (console) multithreaded imaging processing application and deliver it to your clients. You start getting conflicting feedback. Some say that it works perfectly, while others complain that it's slow or sometimes slow. Moreover, different people have vastly different run times, while processing the same dataset on comparable machines. What would be the most likely cause of these issues?
6. Given the following program:

```
for (int var = start; var <= start + 1; var++) {}
```

How does `start` need to be initialized so that the code above runs indefinitely?

2.2 Image processing

1. You are given hyperspectral³ images that are 5000 x 5000 pixels x 50 channels⁴. Each pixel in a channel is 32 bits (4 bytes). You need to process these images. How would you store this image in memory so that you either
 - (a) maximize speed on operations requiring access to spatial information
 - (b) maximize speed on operations requiring access to color information

Provide an explanation for your reasoning

2. You are given two hyperspectral images of 5000 x 5000 pixels x 50 channels. One is sparsely populated, while one is densely populated. Provide the data structures that you would use to store this information.
3. Define the relation between blurring and convolution.

3 Assignments

- You have a few POJOs similar to this one, that hold different types of metadata of an image.

```
public class ImageMetadata {  
  
    private int bands;  
    private String description;  
    private Number height;  
    private Number predictor;  
    private Number rowsPerStrip;  
    private String software;  
    private Number width;  
  
    public ImageMetadata() {  
    }  
  
    public int getBands() {  
        return bands;  
    }  
  
    public String getDescription() {  
        return description;  
    }  
  
    public Number getHeight() {
```

³Hyperspectral images contain many more color channels than the usual three channels for RGB

⁴Channel = layer (or color)

```

        return height;
    }

    public Number getPredictor() {
        return predictor;
    }

    public Number getRowsPerStrip() {
        return rowsPerStrip;
    }

    public String getSoftware() {
        return software;
    }

    public Number getWidth() {
        return width;
    }
}

```

Provide a solution that would allow you to present the different POJOs in the following format:

bands	Number of bands in the image	3
description	Image details / description	"This is a test image"
height	Image height	600
predictor	Predictor applied to the image before encoding is applied	2
rowsPerStrip	Number of rows per strip, in a stripped image	50
software	Software used to produce the image	SpectRAW
width	Image width	800

You need to provide an easy (elegant) way of attaching a description to each property.

Please note that the solution you present can modify the contents of the POJOs.

- You are given two sets of images, coming from two different cameras, mounted on a stereoscopic rig, placed on a drone/quadcopter. One of the cameras produces photos in the visible spectrum (400-650nm), while the other produces photos in the near-infrared (650-900nm). Given that the cameras are triggered simultaneously, provide a Java implementation of an algorithm that can do pairwise alignment of the VIS and NIR images.

Input: one VIS image and one NIR image.

Output:

- an error message if the images cannot be aligned
- an image containing two channels, corresponding to the aligned VIS and NIR images

Images can be found [here](#)

- Your task is to design the class diagram for the following project requirements:
 - For each crop field you have 3 known groups of statistics, each consisting of two real values (e.g. group 1 consists of planted area and total area, group 2 consists of infected crop area and total crop area, group 3 is water surface and total area). You are expected to add two new statistics groups in the near future, which will consist of similar pairs of real values. There have been rumors that in 6 months from now each statistic will include a third real value.
 - Fields are grouped into farms and we need the same statistics for a farm that includes the aggregated results from each comprising field (e.g. if the farm has a field with 2 ha of planted area from 3 ha total area and another field of 5 ha of planted area from 7 ha, the farm has 7 ha of planted area from a total of 10 ha).

- The grouping of fields into farms is provided as JSON files with the following structure:

```
{ "farmName": "myFarm1",  
  "fields": [ "field4324",  
              "field1954",  
              "field1864" ] }
```

- The statistics are provided in JSON files with the following structure

```
{ "statistic": "plantedArea",  
  "covered": 5,  
  "total": 10 }
```

- You need to coordinate with the data science team who produces the statistics on how to present the JSON files for your chosen solution.
- The agronomist who is using the system sometimes wants to see the results for only one statistic, but most of times for multiple of them. Also he sometimes uses information to compare two fields in the same farm or the statistics of different farms.