

Practical Semester Report

Lucas Schmidt

November 7, 2017

bla bla

Contents

Part I

Introduction

1 The Functional Genomics Center Zurich

The Functional Genomics Center Zurich (FGCZ) is a department of the ETH Zurich specialized in research and training. The department was founded in 2002 and is since then an expert in support for Omics research supported by the latest technologies. The FGCZ collaborates with Zurich Life Science research community. There are about 40 staff members employed, about six members are from the subject area of computer science and about ten of them are specialized in the area of bioinformatics. The task of the FGCZ is in general a research and training facility with several departments. Therefore the FGCZ is composed of six specialized areas (Genomics/Transcriptomics, Proteomics & Protein Analysis, Metabolomics & Biophysics, Data Integration & IT, Research Collaborations and Governance & Management).

http://www.fgcz.ch/the-center/_jcr_content/par/fullwidthimage/image.imageformat.lightbox.1160524864.png/

- founded 2002
- task of FGCZ
- structure 40 staff members, 6 computer scientists, 10 bioinformaticians

2 Environment

The Environment includes 20 mass spectrometry devices and 5 DNA sequencer. There are more than 1000 active nodes of Microsoft Windows, Linux and Mac OSX. The data management system is called b-fabric and there is about 500TB of storage. For my project I used a development environment on Linux using Java 8.

- 20 Mass spec devices, 5 sequencer
- 1000 active nodes windows linux mac os x
- firewall uzh
- storage size 500TB (TODO(MOS))
- data management system called b-fabric [?] ...
- dev on Linux Box using Java 8

3 My Project Tasks

Implementation of deisotoper see next chapter
in the follwong sections

Part II

Deisotoping

4 Problem Definition

4.1 Chemical view - What is an isotope pattern?

4.2 Modeling isotopic pattern

4.3 Deconvoluting isotopic pattern

4.4 General deisotoper algorithm

4.5 Characteristics of Deisotoping

map	statistical value	#regions	area err	meta	pop size	mutrate	#gen	compute hardware	cores	proc time	unit	Fig
US state level	#electors	49	0.36	GA	unknown	unknown	10	Intel Xeon CPU @ 1.5 GHz	1	≈ 1	min	??
US state level	area	50	0.17	-	-	-	-	-	1	$\ll 1$	sec	??
8x8 checkerboard	1:4	64	0.27	GA	256	0.25	533	3 GHz Intel Core i7	4	≈ 1	min	??
8x8 checkerboard	1:4	64	0.27	GRASP	NA	NA	NA	3 GHz Intel Core i7	4	≈ 1	min	??
US state level	population 1977	48	0.41	GA	300	0.25	11	3 GHz Intel Core i7	4	5	sec	??
US state level	murder 1977	48	0.37	GA	300	0.25	20	3 GHz Intel Core i7	4	10	sec	??
US state level	income 1977	48	0.29	GA	300	0.25	30	3 GHz Intel Core i7	4	16	sec	??
US state level	illiteracy 1977	48	0.40	GA	300	0.25	30	3 GHz Intel Core i7	4	16	sec	??
US California	population 2010	48	0.62	GA	240	0.25	27	3 GHz Intel Core i7	4	10	sec	??
US Colorado	population 2010	64	0.68	GA	320	0.25	102	3 GHz Intel Core i7	4	74	sec	??
US Florida	population 2010	68	0.44	GA	240	0.25	49	3 GHz Intel Core i7	4	39	sec	??
US New Jersey	population 2010	21	0.38	GA	105	0.25	30	3 GHz Intel Core i7	4	4	sec	??
US New York	population 2010	62	0.62	GA	310	0.25	148	3 GHz Intel Core i7	4	106	sec	??
Switzerland	population 2010	2300	0.59	GA	1280	0.35	300	Intel Xeon CPU E5-2698 v3 @ 2.30GHz	64	3	days	??
Swiss SBB railway	passenger frequency	724	NA	GA	1000	0.25	1000	Intel Xeon CPU E5-2698 v3 @ 2.30GHz	64	2	days	??
UK	electorates	370	0.57	GA	1200	0.25	4000	Intel Xeon CPU E5-2698 v3 @ 2.30GHz	64	3	days	??

Table 4.1: bla bla

5 Related Work

Available Deisotoping software

6 Implementation

Java8 gSON, jGraphT,
make and maven
Rpackage, ...
Test Driven Development (TDD)
Unit Test

7 Evaluation

7.1 Datasets for comparison / Sample data

hela...

8 Conclusion

```
> library(deisotoper)
> Test
```

Part III

System Administration

A R package documentation

B LOG of the 95 days