SAVITRIBAI PHULE PUNE UNIVERSITY

A PRELIMINARY PROJECT REPORT ON

BE PROJECT TITLE

SUBMITTED TOWARDS THE
PARTIAL FULFILLMENT OF THE REQUIREMENTS OF

BACHELOR OF ENGINEERING (Computer Engineering)

BY

Rathin Desai	Exam No:
Shubhankar Deshpande	Exam No:
Archit Sakhadeo	Exam No:
Shadab Shaikh	Exam No:

Under The Guidance of

Prof. Y.A. Handge

DEPARTMENT OF COMPUTER ENGINEERING
Pune Institute of Computer Technology

College Name
DEPARTMENT OF COMPUTER ENGINEERING

CERTIFICATE

This is to certify that the Project Entitled

BE PROJECT TITLE

Submitted by

Rathin Desai	Exam No:
Shubhankar Deshpande	Exam No:
Archit Sakhadeo	Exam No:
Shadab Shaikh	Exam No:

is a bonafide work carried out by Students under the supervision of Prof. Y.A. Handge and it is submitted towards the partial fulfillment of the requirement of Bachelor of Engineering (Computer Engineering) Project.

Prof. Y.A. Handge	Prof. R.B. Ingle
Internal Guide	H.O.D
Dept. of Computer Engg.	Dept. of Computer Engg.

Abstract

Radio interferometric image generation is a tedious task, with numerous weeks of manual effort required to be invested before the synthesis of a single image. The time complexity of manual image synthesis can be attributed to multiple factors including flagging of data, and the inherently interactive nature of existing tools used for high resolution astronomical imaging. This situation is expected to further worsen with the development of wide bandwidth upgrades for the GMRT, which will lead to increased data throughput, and hence increased processing time (assuming at least a linear relation between data hypercube sizes and time for processing). A collection of AIPS based software for source peeling and atmospheric modeling includes direction-dependent and direction-independent calibration, and has been shown to work well on a subset of data collected by the GMRT at different frequencies. Through bootstrapping specific features of this tool, and integrating with a custom high performance computing cluster that we have designed and implemented, we convert raw correlated radio interferemotric data from the GMRT, to synthesize primary beam corrected images in a massively parallel automated manner. Our primary implementation has shown to be robust, and has been tested with over 2000 Gigabytes of data, to synthesize 1000+ images reducing the processing time from 6+ weeks to a mere 10-12 hours.

Acknowledgments

It gives us great pleasure in presenting the preliminary project report on ***GMRT ARCHIVAL UTILILTY FOR DATA ANALYSIS***.

I would like to take this opportunity to thank my internal guide ***Prof. Y.A. Handge*** *for giving me all the help and guidance I needed. I am really grateful to them for their kind support. Their valuable suggestions were very helpful.*

I am also grateful to ***Prof. R.B. Ingle****, Head of Computer Engineering Department, CollegeName for his indispensable support, suggestions.*

In the end our special thanks to ***Other Person Name*** *for providing various resources such as laboratory with all needed software platforms, continuous Internet connection, for Our Project.*

Rathin Desai
Shubhankar Deshpande
Archit Sakhadeo
Shadab Shaikh
(B.E. Computer Engineering)

INDEX

List of Figures

List of Tables

CHAPTER 1

SYNOPSIS

1.1 PROJECT TITLE

GMRT Archival Utility for Data Analysis

1.2 PROJECT OPTION

Research Institute Sponsored

1.3 INTERNAL GUIDE

Prof. Y.A. Handge

1.4 SPONSORSHIP AND EXTERNAL GUIDE

National Center for Radio Astrophysics, Tata Institute of Fundamental Research (NCRA-TIFR)
Dr.Yogesh Wadadekar & Dr.C.H. Ishwara Chandra

1.5 TECHNICAL KEYWORDS (AS PER ACM KEYWORDS)

Please note ACM Keywords can be found : http://www.acm.org/about/class/ccs98-html
Example is given as

1. C. Computer Systems Organization
   (a) C.2 COMPUTER-COMMUNICATION NETWORKS
      i. C.2.4 Distributed Systems
         A. Client/server
         B. Distributed applications
         C. Distributed databases
         D. Network operating systems
         E. Distributed file systems
         F. Security and reliability issues in distributed applications

1.6 PROBLEM STATEMENT

Define Problem Statement

1.7 ABSTRACT

- Abstract (10 to 15 lines)

1.8 GOALS AND OBJECTIVES

- Objectives

1.9 RELEVANT MATHEMATICS ASSOCIATED WITH THE PROJECT

System Description:

- Input:
- Output:
- Identify data structures, classes, divide and conquer strategies to exploit distributed/parallel/concurrent processing, constraints.
- Functions : Identify Objects, Morphisms, Overloading in functions, Functional relations
- Mathematical formulation if possible
- Success Conditions:
- Failure Conditions:

1.10 NAMES OF CONFERENCES / JOURNALS WHERE PAPERS CAN BE PUBLISHED

- IEEE/ACM Conference/Journal 1
- Conferences/workshops in IITs

- Central Universities or SPPU Conferences
- IEEE/ACM Conference/Journal 2

1.11 REVIEW OF CONFERENCE/JOURNAL PAPERS SUPPORTING PROJECT IDEA

Atleast 10 papers + White papers or web references
Brief literature survey [Description containing important description of at least 10 papers

1.12 PLAN OF PROJECT EXECUTION

Using planner or alike project management tool.

CHAPTER 2

TECHNICAL KEYWORDS

2.1 AREA OF PROJECT

Project Area

2.2 TECHNICAL KEYWORDS

Please note ACM Keywords can be found : http://www.acm.org/about/class/ccs98-html
Example is given as

1. C. Computer Systems Organization
    - (a) C.2 COMPUTER-COMMUNICATION NETWORKS
        - i. C.2.4 Distributed Systems
            - A. Client/server
            - B. Distributed applications
            - C. Distributed databases
            - D. Network operating systems
            - E. Distributed file systems
            - F. Security and reliability issues in distributed applications

CHAPTER 3

INTRODUCTION

3.1 PROJECT IDEA

- Project Idea

3.2 MOTIVATION OF THE PROJECT

- Motivation of the Project

3.3 LITERATURE SURVEY

- Review of the papers, Description , Mathematical Terms

CHAPTER 4

PROBLEM DEFINITION AND SCOPE

4.1 PROBLEM STATEMENT

Description of Problem

4.1.1 Goals and objectives

Goal and Objectives:

- Overall goals and objectives of software, input and output description with necessary syntax, format etc are described

4.1.2 Statement of scope

- A description of the software with Size of input, bounds on input, input validation, input dependency, i/o state diagram, Major inputs, and outputs are described without regard to implementation detail.
- The scope identifies what the product is and is not, what it will and wont do, what it will and wont contain.

4.2 SOFTWARE CONTEXT

- The business or product line context or application of the software is to be given

4.3 MAJOR CONSTRAINTS

- Any constraints that will impact the manner in which the software is to be specified, designed, implemented or tested are noted here.

4.4 METHODOLOGIES OF PROBLEM SOLVING AND EFFICIENCY ISSUES

- The single problem can be solved by different solutions. This considers the performance parameters for each approach. Thus considers the efficiency issues.

4.5 SCENARIO IN WHICH MULTI-CORE, EMBEDDED AND DISTRIBUTED COMPUTING USED

Explain the scenario in which multi-core, embedded and distributed computing methodology can be applied.

4.6 OUTCOME

- Outcome of the project

4.7 APPLICATIONS

- Applications of Project

4.8 HARDWARE RESOURCES REQUIRED

Sr. No.	Parameter	Minimum Requirement	Justification
1	CPU Speed	2 GHz	Remark Required
2	RAM	3 GB	Remark Required

Table 4.1: Hardware Requirements

4.9 SOFTWARE RESOURCES REQUIRED

Platform :

1. Operating System:
2. IDE:
3. Programming Language

CHAPTER 5

PROJECT PLAN

5.1 PROJECT ESTIMATES

Use Waterfall model and associated streams derived from assignments 1,2, 3, 4 and 5(Annex A and B) for estimation.

5.1.1 Reconciled Estimates

5.1.1.1 Cost Estimate

5.1.1.2 Time Estimates

5.1.2 Project Resources

Project resources [People, Hardware, Software, Tools and other resources] based on Memory Sharing, IPC, and Concurrency derived using appendices to be referred.

5.2 RISK MANAGEMENT W.R.T. NP HARD ANALYSIS

This section discusses Project risks and the approach to managing them.

5.2.1 Risk Identification

For risks identification, review of scope document, requirements specifications and schedule is done. Answers to questionnaire revealed some risks. Each risk is categorized as per the categories mentioned in [1]. Please refer table 5.1 for all the risks. You can refereed following risk identification questionnaire.

1. Have top software and customer managers formally committed to support the project?
2. Are end-users enthusiastically committed to the project and the system/product to be built?
3. Are requirements fully understood by the software engineering team and its customers?
4. Have customers been involved fully in the definition of requirements?
5. Do end-users have realistic expectations?

6. Does the software engineering team have the right mix of skills?
7. Are project requirements stable?
8. Is the number of people on the project team adequate to do the job?
9. Do all customer/user constituencies agree on the importance of the project and on the requirements for the system/product to be built?

5.2.2 Risk Analysis

The risks for the Project can be analyzed within the constraints of time and quality

ID	Risk Description	Probability	Impact		
			Schedule	Quality	Overall
1	Description 1	Low	Low	High	High
2	Description 2	Low	Low	High	High

Table 5.1: Risk Table

Probability	Value	Description
High	Probability of occurrence is	$> 75\%$
Medium	Probability of occurrence is	$26 - 75\%$
Low	Probability of occurrence is	$< 25\%$

Table 5.2: Risk Probability definitions [1]

Impact	Value	Description
Very high	$> 10\%$	Schedule impact or Unacceptable quality
High	$5 - 10\%$	Schedule impact or Some parts of the project have low quality
Medium	$< 5\%$	Schedule impact or Barely noticeable degradation in quality Low Impact on schedule or Quality can be incorporated

Table 5.3: Risk Impact definitions [1]

5.2.3 Overview of Risk Mitigation, Monitoring, Management

Following are the details for each risk.

Risk ID	1
Risk Description	Description 1
Category	Development Environment.
Source	Software requirement Specification document.
Probability	Low
Impact	High
Response	Mitigate
Strategy	Strategy
Risk Status	Occurred

Risk ID	2
Risk Description	Description 2
Category	Requirements
Source	Software Design Specification documentation review.
Probability	Low
Impact	High
Response	Mitigate
Strategy	Better testing will resolve this issue.
Risk Status	Identified

5.3 PROJECT SCHEDULE

5.3.1 Project task set

Major Tasks in the Project stages are:

- Task 1:
- Task 2:
- Task 3:

Risk ID	3
Risk Description	Description 3
Category	Technology
Source	This was identified during early development and testing.
Probability	Low
Impact	Very High
Response	Accept
Strategy	Example Running Service Registry behind proxy balancer
Risk Status	Identified

- Task 4:
- Task 5:

5.3.2 Task network

Project tasks and their dependencies are noted in this diagrammatic form.

5.3.3 Timeline Chart

A project timeline chart is presented. This may include a time line for the entire project. Above points should be covered in Project Planner as Annex C and you can mention here Please refer Annex C for the planner

5.4 TEAM ORGANIZATION

The manner in which staff is organized and the mechanisms for reporting are noted.

5.4.1 Team structure

The team structure for the project is identified. Roles are defined.

5.4.2 Management reporting and communication

Mechanisms for progress reporting and inter/intra team communication are identified as per assessment sheet and lab time table.

CHAPTER 6

SOFTWARE REQUIREMENT SPECIFICATION (SRS IS TO BE PREPARED USING RELEVANT MATHEMATICS DERIVED AND SOFTWARE ENGG. INDICATORS IN ANNEX A AND B)

6.1 INTRODUCTION

6.1.1 Purpose and Scope of Document

The purpose of SRS and what it covers is to be stated

6.1.2 Overview of responsibilities of Developer

What all activities carried out by developer?

6.2 USAGE SCENARIO

This section provides various usage scenarios for the system to be developed.

6.2.1 User profiles

The profiles of all user categories are described here.(Actors and their Description)

6.2.2 Use-cases

All use-cases for the software are presented. Description of all main Use cases using use case template is to be provided.

Sr No.	Use Case	Description	Actors	Assumptions
1	Use Case 1	Description	Actors	Assumption

Table 6.1: Use Cases

6.2.3 Use Case View

Use Case Diagram. Example is given below

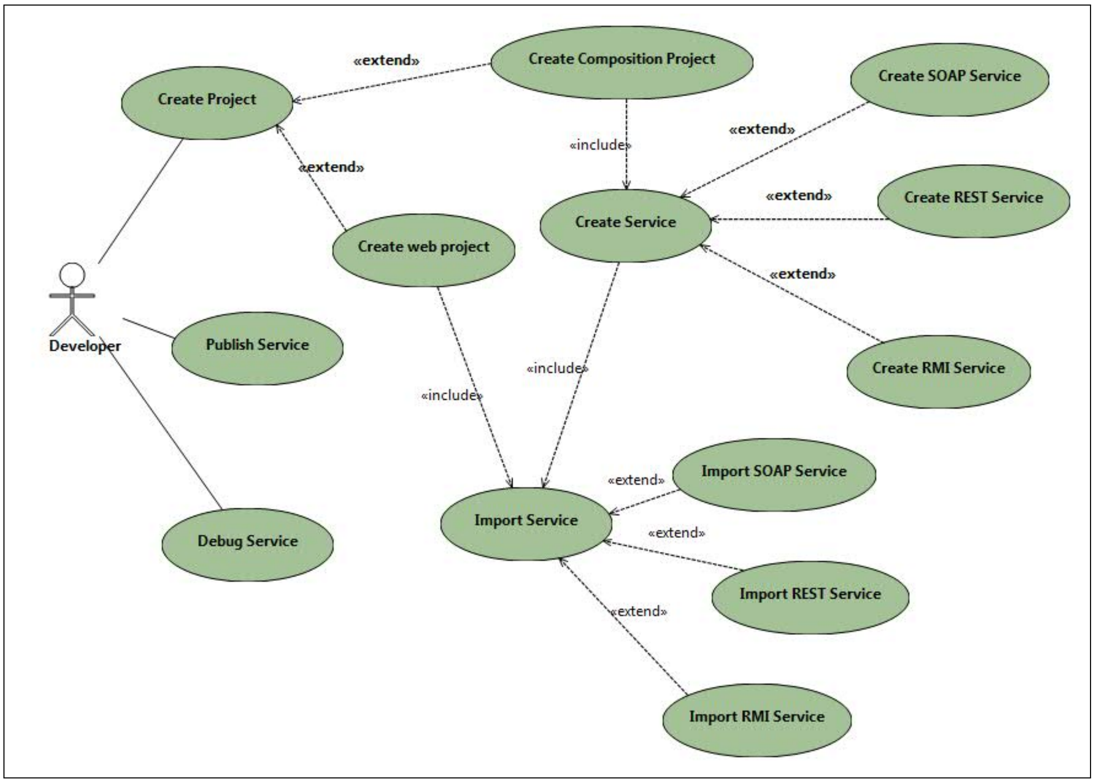

Figure 6.1: Use case diagram

6.3 DATA MODEL AND DESCRIPTION

6.3.1 Data Description

Data objects that will be managed/manipulated by the software are described in this section. The database entities or files or data structures required to be described. For data objects details can be given as below

6.3.2 Data objects and Relationships

Data objects and their major attributes and relationships among data objects are described using an ERD- like form.

6.4 FUNCTIONAL MODEL AND DESCRIPTION

A description of each major software function, along with data flow (structured analysis) or class hierarchy (Analysis Class diagram with class description for object oriented system) is presented.

6.4.1 Data Flow Diagram

6.4.1.1 Level 0 Data Flow Diagram

6.4.1.2 Level 1 Data Flow Diagram

6.4.2 Description of functions

A description of each software function is presented. A processing narrative for function n is presented.(Steps)/ Activity Diagrams. For Example Refer 6.2

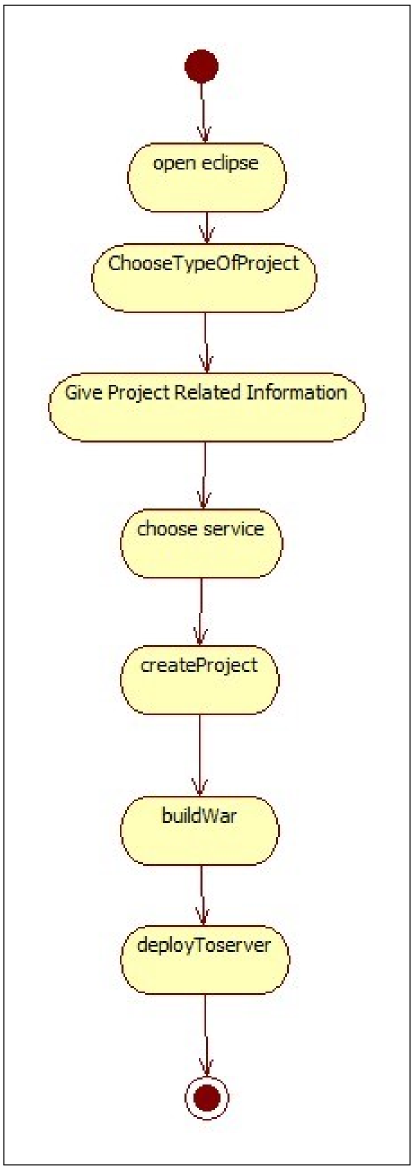

Figure 6.2: Activity diagram

6.4.3 Activity Diagram:

- The Activity diagram represents the steps taken.

6.4.4 Non Functional Requirements:

- Interface Requirements
- Performance Requirements
- Software quality attributes such as availability [related to Reliability], modifiability [includes portability, reusability, scalability] , performance, security, testability and usability[includes self adaptability and user adaptability]

6.4.5 State Diagram:

State Transition Diagram
Fig.6.3 example shows the state transition diagram of Cloud SDK. The states are represented in ovals and state of system gets changed when certain events occur. The transitions from one state to the other are represented by arrows. The Figure shows important states and events that occur while creating new project.

6.4.6 Design Constraints

Any design constraints that will impact the subsystem are noted.

6.4.7 Software Interface Description

The software interface(s)to the outside world is(are) described. The requirements for interfaces to other devices/systems/networks/human are stated.

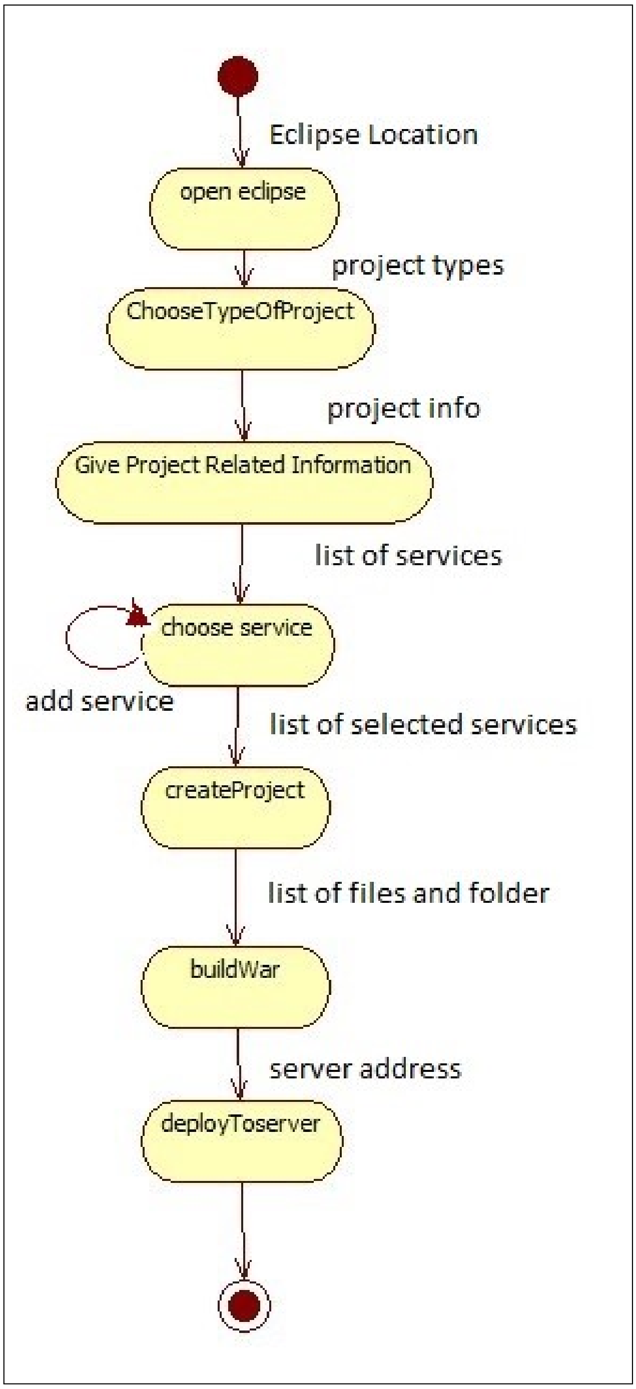

Figure 6.3: State transition diagram

CHAPTER 7

DETAILED DESIGN DOCUMENT USING APPENDIX A AND B

7.1 INTRODUCTION

This document specifies the design that is used to solve the problem of Product.

7.2 ARCHITECTURAL DESIGN

A description of the program architecture is presented. Subsystem design or Block diagram,Package Diagram,Deployment diagram with description is to be presented.

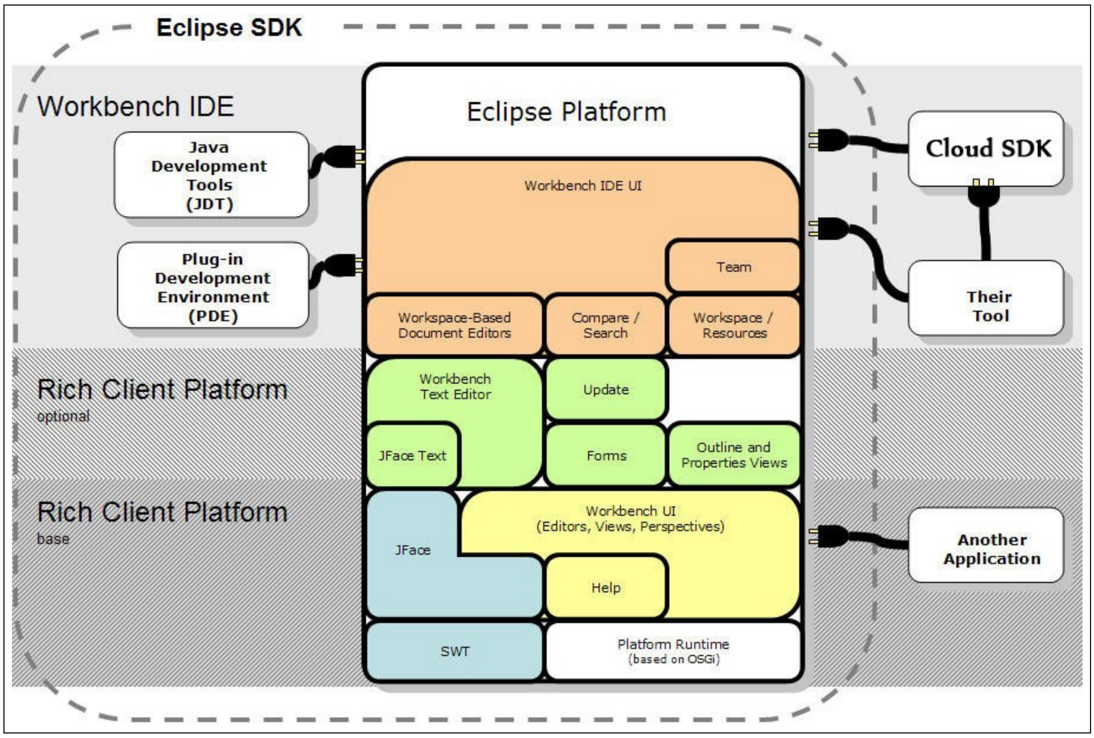

Figure 7.1: Architecture diagram

7.3 DATA DESIGN (USING APPENDICES A AND B)

A description of all data structures including internal, global, and temporary data structures, database design (tables), file formats.

7.3.1 Internal software data structure

Data structures that are passed among components the software are described.

7.3.2 Global data structure

Data structured that are available to major portions of the architecture are described.

7.3.3 Temporary data structure

Files created for interim use are described.

7.3.4 Database description

Database(s) / Files created/used as part of the application is(are) described.

7.4 COMPOENT DESIGN

Class diagrams, Interaction Diagrams, Algorithms. Description of each component description required.

7.4.1 Class Diagram

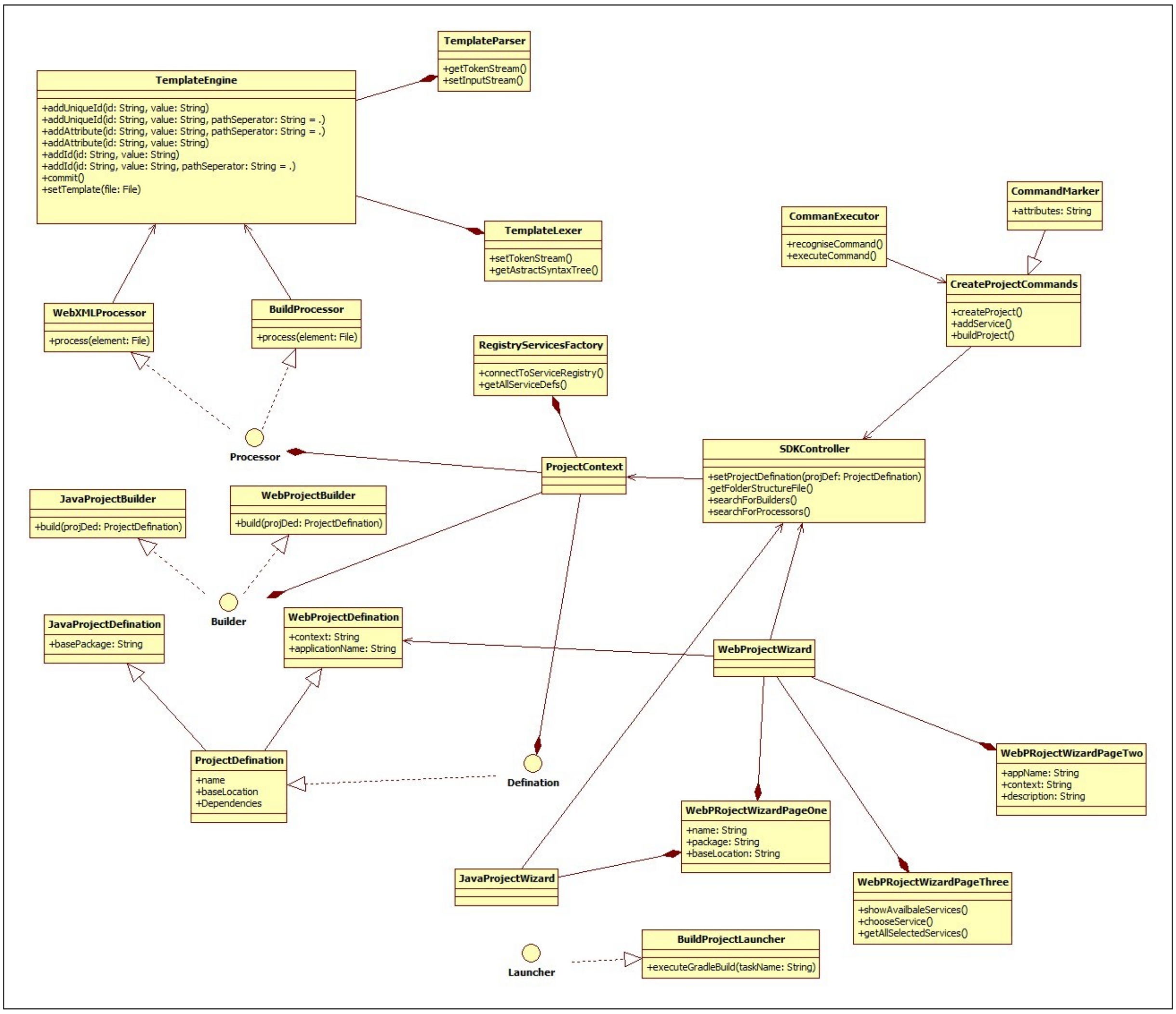

Figure 7.2: Class Diagram

CHAPTER 8

SUMMARY AND CONCLUSION

Write one page summary and conclusion

CHAPTER 9

REFERENCES

[1] R. S. Pressman, *Software Engineering (3rd Ed.): A Practitioner's Approach*. New York, NY, USA: McGraw-Hill, Inc., 1992.

[2] P. Kulkarni, *Knowledge Innovation Strategy*. Pune: Bloomsbury Publication, 2015.

[3] P. Sinha, *Electronic Health Record*. IEEE Press Wiley.

[4] McKinsey, "Big data: The next frontier for innovation, competition, and productivity," tech. rep.

[5] "Digital competition:http://www.mckinsey.com/insights."

[6] "Government website to support entreprenaurs:http://msme.gov.in/mob/home.aspx."

ANNEXURE A

LABORATORY ASSIGNMENTS ON PROJECT ANALYSIS OF ALGORITHMIC DESIGN

- To develop the problem under consideration and justify feasibilty using concepts of knowledge canvas and IDEA Matrix.

  Refer [2] for IDEA Matrix and Knowledge canvas model. Case studies are given in this book. IDEA Matrix is represented in the following form. Knowledge canvas represents about identification of opportunity for product. Feasibility is represented w.r.t. business perspective.

I	D	E	A
Increase	Drive	Educate	Accelerate
Improve	Deliver	Evaluate	Associate
Ignore	Decrease	Eliminate	Avoid

Table A.1: IDEA Matrix

- Project problem statement feasibility assessment using NP-Hard, NP-Complete or satisfy ability issues using modern algebra and/or relevant mathematical models.

- input x,output y, y=f(x)

ANNEXURE B

LABORATORY ASSIGNMENTS ON PROJECT QUALITY AND RELIABILITY TESTING OF PROJECT DESIGN

It should include assignments such as

- Use of divide and conquer strategies to exploit distributed/parallel/concurrent processing of the above to identify object, morphisms, overloading in functions (if any), and functional relations and any other dependencies (as per requirements). It can include Venn diagram, state diagram, function relations, i/o relations; use this to derive objects, morphism, overloading

- Use of above to draw functional dependency graphs and relevant Software modeling methods, techniques including UML diagrams or other necessities using appropriate tools.

- Testing of project problem statement using generated test data (using mathematical models, GUI, Function testing principles, if any) selection and appropriate use of testing tools, testing of UML diagram's reliability. Write also test cases [Black box testing] for each identified functions. You can use Mathematica or equivalent open source tool for generating test data.

- Additional assignments by the guide. If project type as Entreprenaur, Refer [3],[4],[5], [6]

ANNEXURE C

PROJECT PLANNER

Using planner or alike project management tool.

ANNEXURE D

REVIEWERS COMMENTS OF PAPER SUBMITTED

(At-least one technical paper must be submitted in Term-I on the project design in the conferences/workshops in IITs, Central Universities or UoP Conferences or equivalent International Conferences Sponsored by IEEE/ACM)

1. Paper Title:
2. Name of the Conference/Journal where paper submitted :
3. Paper accepted/rejected :
4. Review comments by reviewer :
5. Corrective actions if any :

ANNEXURE E

PLAGIARISM REPORT

Plagiarism report