20 October, 2014



Team Number: 4 Consultation Day/Hour: Tuesday, 1pm

Team Name: Team 4

Team Members Information:

Group PKB

Saloni Kaur A0084053L [a0084053@nus.edu.sg](mailto:a0084053@nus.edu.sg)

M I Azima A0085594N [a0085594@nus.edu.sg](mailto:a0085594@nus.edu.sg)

Group PQL

Saima Mahmood A0084176Y [a0084176@nus.edu.sg](mailto:a0084176@nus.edu.sg)

Nguyen Trong Son A0088441 [A0088441@nus.edu.sg](mailto:A0088441@nus.edu.sg)

Vu Phuc Tho A0090585X [A0090585@nus.edu.sg](mailto:A0090585@nus.edu.sg)

CS3202 Software Engineering Project

Iteration 3 Report

# Iteration Overview

## Scope of SPA Implemented

In this second iteration we have followed the suggested implementation of the SPA according to the assignment document. We have extended the functionality of the PKB to include the Affects relationship. The QP has been extended to deal with tuple results. The Query Evaluator has been optimized as well.

## Achievements & Problems

The main achievement for this iteration would have to be the fact that we managed to implement the requirements expected from this iteration. We also implemented a design extractor to ease the functionality from the parser. Furthermore we also decided to change the underlying data structure used to construct all of the PKB tables for the various relations. This will be discussed further in depth in section 4.

From iteration two we picked up on some problems of our system, with the building of the next and affects relations. This, will be further elaborated on in section 8.

# Project Plans

The way that we have decided to break down the project into tasks is shown in tables 1 and 2 below.

## Plan for Whole Project

The breakdown of tasks reflected in table 1 is tentative and will be subjected to change as we move along the various iterations. The subsequent changes, if any, shall be reflected in the reports of the corresponding iteration.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Iteration 1** | | | | **Iteration 2** | | | **Iteration 3** | | |
| **Team Member** | PKB | Parser | Query Processer | Report | PKB | Query Processer | Report | Affects Relationship | Tuple Results | Report |
| Azima | \* |  |  | \* | \* |  | \* |  |  | \* |
| Saima |  |  | \* | \* |  | \* | \* |  |  | \* |
| Saloni | \* |  |  | \* | \* |  | \* |  |  | \* |
| Sean |  | \* |  | \* |  | \* | \* |  |  | \* |
| Tho |  |  | \* |  |  | \* | \* |  |  | \* |

Table 1: Whole Project Tasks Breakdown

## Plan for Iteration 3

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Team Member** | **Testing** | **Writing Test Cases** | **Revamp of Relationships Data Structure** | **Affects** | **Working on QP** | **Fixing Issues with QE** | **Report** |
| Azima | \* | \* |  |  |  |  | \* |
| Saima | \* |  |  |  | \* |  | \* |
| Saloni | \* | \* |  |  |  |  | \* |
| Sean | \* |  | \* | \* |  |  |  |
| Tho | \* |  |  |  | \* | \* |  |

Table 2: Iteration 3 Work Distribution

# UML Diagrams

\*SAME AS BEFORE\*

# Design Decisions

In this section we will discuss the data structure we have used in representing the relationship tables in the PKB. The implementation of Affects/Affects Star and Next Star will also be discussed. The implementation of the Design Extractor will also be discussed.

## PKB Tables’ Data Structure

In this iteration of our SPA development, we have decided to revamp the way in which we represent the PKB tables. It has to be noted that previously in iteration 2, the tables were implemented via 2D vectors storing Boolean values. Fig. 1 shows a representation of the *Follow* table as was previously implemented.

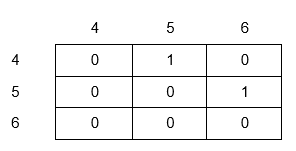


Figure 1: Previous Representation of Follows Table

However we realized that when the size of the table would have to be extended the time complexity to resize the table would be O(n^2). As such we decided to change this to an implementation where we do not have to resize the tables.

Now we have decided to replace all of the 2D vector representations with a MapTable, for all relationship tables, and a ListTable for the remaining tables such as the Symbol Table. The ListTable has been implemented by a single vector, while the MapTable consists of 4 Map attributes. All of this is displayed in fig 2.

Figure 2: ADT used for PKB

As can be seen from the figure, the 4 attributes for MapTable will store the various relationship information in the form of a key-value of a map. To capture the 2 way relationship of any entity encountered, we have implemented 2 maps, keyValue and valueKey, to keep track of the forward and backward direction relationship. keyValueStar and valueKeyStar similarly keeps track of the corresponding star relationships.

To give an example, the implementation of the Next Table could look like as follows.

MapTable <int> NextTable;

keyValue (Map of type <int, vector<int>>)

Possible value pairs: (1,2), (2,3), (3, 4), (4, 5)

valueKey (Map of type <int, vector<int>>)

Possible value pairs: (2,1), (3,2), (4, 3), (5, 4)

keyValueStar (Map of type <int, vector<int>>)

Possible value pairs: (1, [2, 3, 4, 5])

The reason for keeping track of the backward direction relationship is to ease the computation of the queries when the 1st or 2nd argument is unknown. Based on the unknown value, the PKB will access the relevant maps.

## Affects/ Affects Star

For the implementation the Affects relationship, and the star relationship by extension, we have added methods to determine Affects in the PKB. Since affects is calculated on the fly during the query evaluation, Affects does not have a table per say. Based on a given query, the QE will call a method in the PKB to determine the Affects/Affects star relationship. Then in the PKB, the corresponding method will do the following:

1. First of all check if the two arguments are in the same procedure. This will be done by a breadth first search in the CFG paths of the 2 arguments. The CFG is basically the Next Table.
2. The next thing to occur is that all possible CFG paths will be discovered.
3. After that, for every CFG path the method will check if between the start and end program line of this CFG path, a variable is not modified in any assignment statement or procedure call
4. A list of results will then be returned to the QE to be returned to the main controller.

Affects star also works similarly by calling the relevant affects methods in the PKB in a loop, until all the possible answers that satisfy the query are found.

## Design Extractor

As part of the earlier iterations of the SPA we did not deem it necessary to implement a Design Extractor, DE, since the Next, Affects relationships were not to be implemented. However as part of this iteration, it was necessary to have included a DE to ease the functionality of the Parser. The flow of the SPA, with the inclusion of the DE, now works in the manner as shown in fig 3.

2.

1.

Figure 3: Parser & DE Program Flow

As seen from fig. 3, we can see how the DE takes up more of the complicated functionality of implementing the relationships in the PKB. In step 2, the DE also completes the Modifies and Uses tables for procedure calls and Modifies/ Uses statements.

# Coding Standards & Experiences

\*SAME AS BEFORE\*

# Query Processing

In this section we will go through how queries are evaluated. For the discussion below we will be referring to the sample query, shown below, to illustrate the process.

stmt s1, s2, s3;

Select s1 such that Parent\*(s1, s2) pattern s3(\_, \_“x+y\*2”\_)

## Query Validation

The query validating process is handled by the QueryPreprocessor (QPP). The QPP will read and validate data from the query file, and store all necessary information for query evaluation in the Query Representator (QR).

During validation, the QPP checks each query line by line, following these rules:

- On finding a line containing a declaration (e.g. stmt s;), QPP will send this part to the preprocessDeclaration() method.

- On finding a line containing a query part (e.g. Select s such that Follows(s, 1)), QPP will send this part to preprocessQueryPart() method.

- While the preprocessQueryPart() method runs, if the QPP finds a new clause of query, it will call the corresponding method preprocessClause() (e.g. preprocessSuchThatCondtion() ) for this clause. Currently, we have provided the methods for “*such that*” clause and “*pattern*” clause.

After validating the query, if no error is found, the QPP will save validated data in the QR. The QR will contain the query tree. An example of what the query tree, based on the sample query, will look like is shown below.

## Query Evaluation

The query evaluation process is carried out by the QE. Using the stored data from the QR and PKB, the QE will try to find values that satisfy the query’s conditions, and send the list of result values to the AutoTester.

During the evaluation of a query, firstly, the QE will create an empty table for the result. After that, it will check all of the query’s conditions one by one, attempting to update the list with values of parameters that satisfy the conditions. If it meets any condition that no value can satisfy, the QE will hang up the process and return an empty list as the result of the query.

After the process finishes, with all conditions being satisfied, the QE will take the result value and save it to the result list. This is the list that will be output to the user.

The sample query shown below will be used as an example to explain the query evaluation process

### Manage the temporary results

|  |
| --- |
| S1 |
| 1 |
| 2 |
| 3 |

Before evaluating a query, the QE creates an empty table. Keeping the above sample query in mind, the updating of this table is illustrated below.

2. Add s1 to table

1. Encounters Parent\* Clause

|  |
| --- |
| \*Empty table\* |

3. For each value of s1, adds the value of s2. Remove the old value of s1.

|  |  |  |
| --- | --- | --- |
| S1 | S2 | S3 |
| ~~1~~ | ~~5~~ |  |
| ~~1~~ | ~~6~~ |  |
| 1 | 6 | 3 |
| 1 | 6 | 4 |

|  |  |
| --- | --- |
| S1 | S2 |
| ~~1~~ |  |
| ~~2~~ |  |
| ~~3~~ |  |
| 1 | 5 |
| 1 | 6 |

4. Ended evaluating Parent\* clause. Move on to Pattern clause

5. Add s3 to the table and update values

6. Final result returned

During processing a query’s condition, if a new value is found, the QE will update the list of values, by adding this new row. The QE will remove the old row and use the updated list to continue the evaluation. The list continues to be updated until there is no more condition left or a condition which cannot be satisfied appears.

In this way there is no need to consider merging results for the various clauses for evaluation. This updating of results table saves us from this hassle.

Finally at the end the result node will be read, from the query tree, to determine which result to return back to the QP.

### Tuple Results

Based on the example shown above for managing the temporary results, it is quite obvious how the QE can also manage returning tuple results. In the event that the result node of the query tree points to returning more than one variable, the QE will just select the corresponding pairs of results from the final result table, as shown above.

### Optimize the evaluation process

To better optimize our query evaluation, we have created a new class called Query Optimizer, QO. The job of this class is to reevaluate the resulting query tree, which has been built by the QPP. In this class 3 conditions are kept in mind to rearrange the query tree. This addition of the QO does not impact how the QE works. This shows good separation of concerns.

Let us now look at the 3 conditions in details.

1. Type of clause. The 3 types of clauses are such that, pattern and with. In this condition, we will arrange the query tree such that the clause with the smallest time complexity will be evaluated first. We have determined that the with clause takes the smallest time complexity to solve, followed by such that and pattern clauses. So the order of importance of clauses are:
   1. With
   2. Such that
   3. Pattern
2. Number of symbols in each clause. It is obvious that the lesser the number of symbols in a clause, the easier it will be to solve. As such we for each clause, the number of symbols will be counted and smaller the number for that clause, we will put it at a higher priority to solve first. The thing to note:
   1. Num of symbols
3. Symbol frequency. Throughout all of the clauses we will determine which symbol is used most frequently. If a clause contains a symbol that is highly used, we will solve that clause first, since it will make the solving of the other clauses easier. The thing to note is:
   1. Freq of each symbol

Now keeping in mind all of the above conditions, we will give each clause a ranking and rearrange the query tree according to this ranking.

Let’s take a look at the query below,

stmt s1, s2, s3;

Select s1 such that Parent\*(s1, s2) pattern s3(\_, \_“x+y\*2”\_) with s1.stmt#=10

Now after optimization, the query shown above would look something like this,

stmt s1, s2, s3;

Select s1 with s1.stmt#=10 such that Parent\*(s1, s2) pattern s3(\_, \_“x+y\*2”\_)

# Testing

## Test Plan

In this iteration we continuously tested each and every component after changes were made. We made the decision that unit testing on each component would be done by the person implementing the component. Integration testing amongst components would be carried out by the people responsible for the specific components. Finally validation testing will be carried out by everyone at various points of progress during development.

In some aspects of the testing we have used assertions. Examples of where they were used will be shown in the figures shown below in the following sub-sections.

## Examples of Test Cases

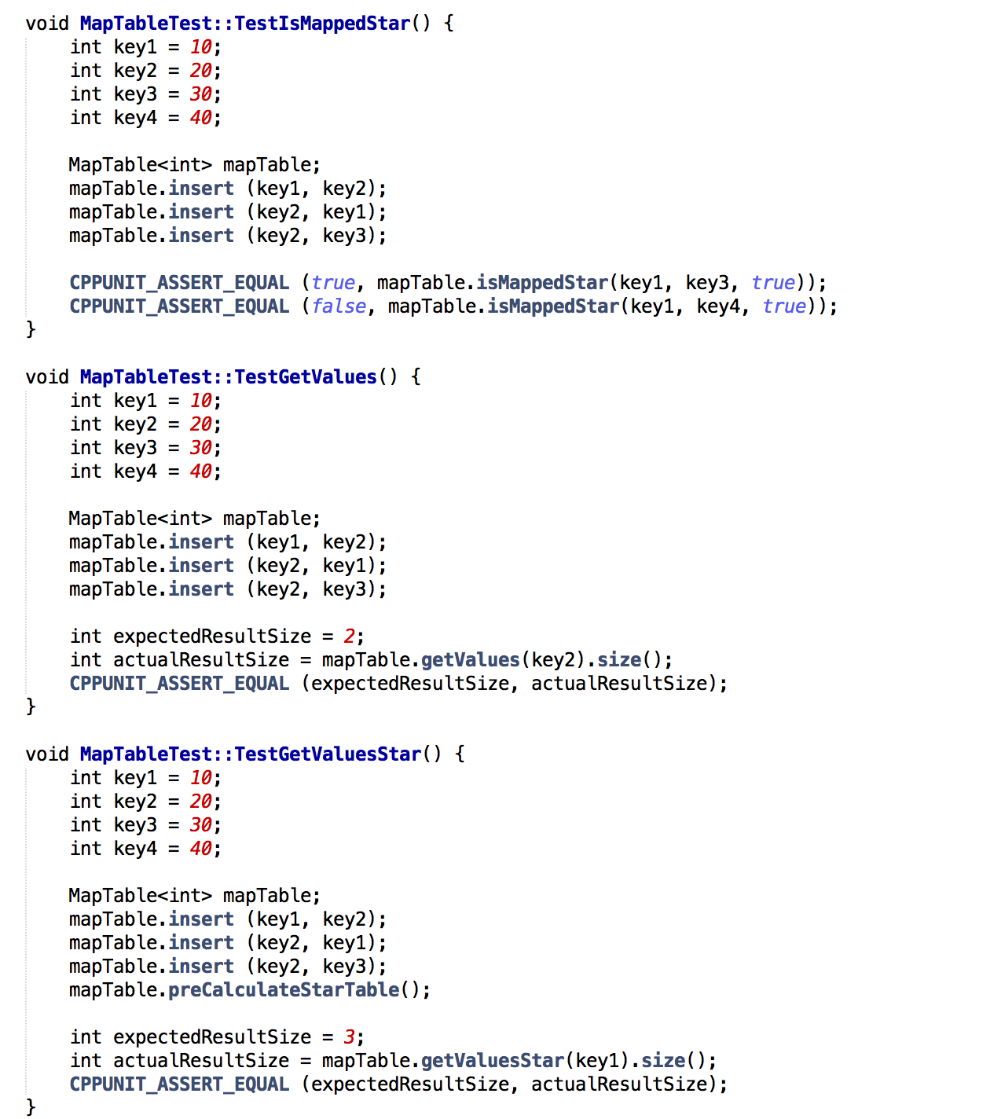
### Unit Testing

#### PKB

Sample 1: Test MapTable

Test Purpose: To Test Map Table

Required Test Inputs: Map Table and CPPUnit test cases. An example is shown below



Expected Test Results: “OK tests” if it passes all CPPUnit asserts

Other Requirements: None

Sample 2: Test List Table

Test Purpose: Test List Table

Required Test Inputs: List Table and CPPUnit test cases. An example is shown below.



Expected Test Results: “OK tests” if it passes all CPPUnit asserts

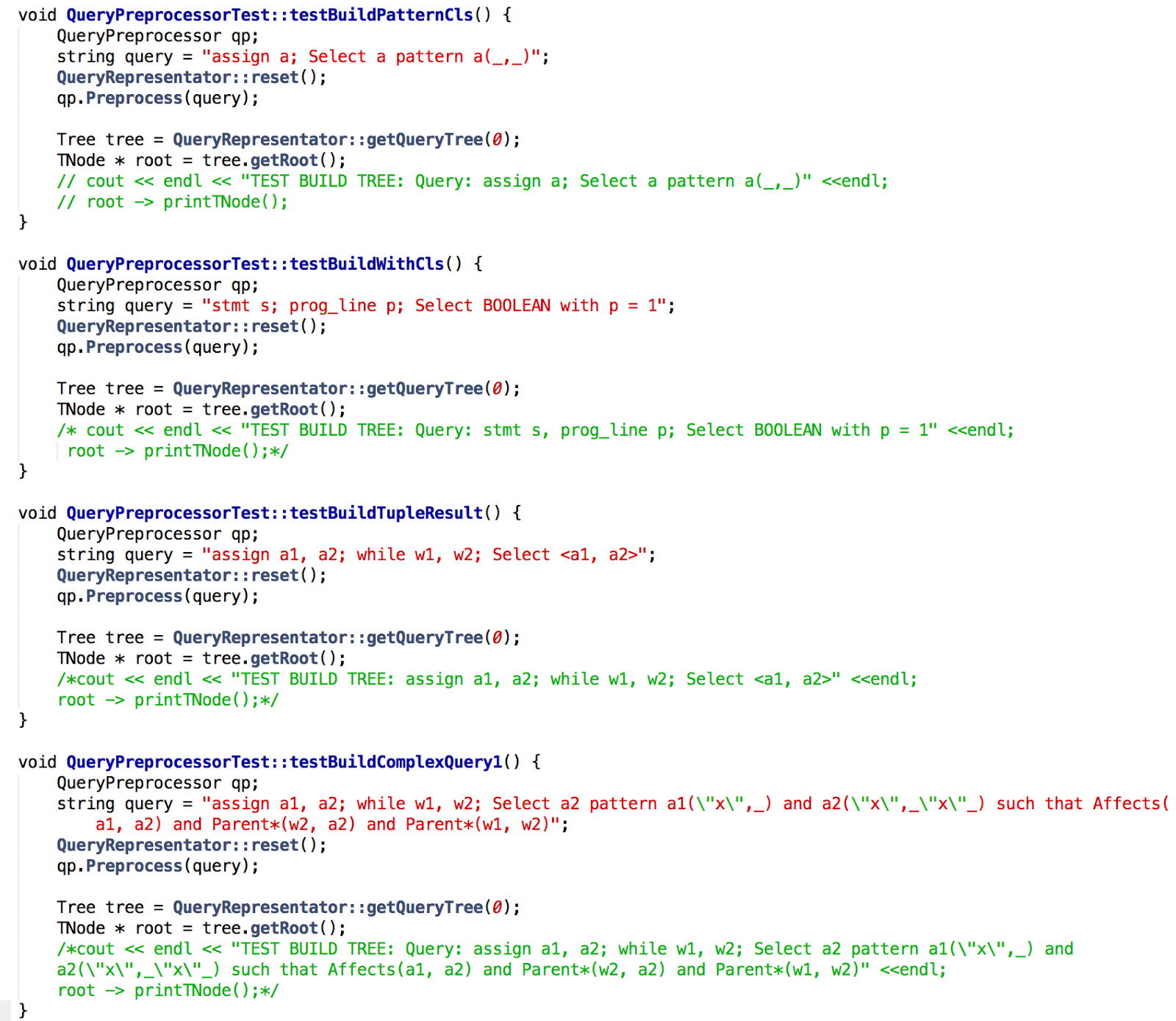
Other Requirements: None

#### PQL

Sample 1: Test QP

Test Purpose: Test Query Processor

Required Test Inputs: Query Preprocessor and PQL Queries. An example is shown below.



Expected Test Results: “OK tests” if it passes all CPPUnit asserts

Other Requirements: None

Sample 2: Test QR

Test Purpose: Test Query Representator

Required Test Inputs: Query Representator and a pre-built QueryTree and SymbolTable. An example is shown below.



Expected Test Results: “OK tests” if it passes all CPPUnit asserts

Any other requirements: None

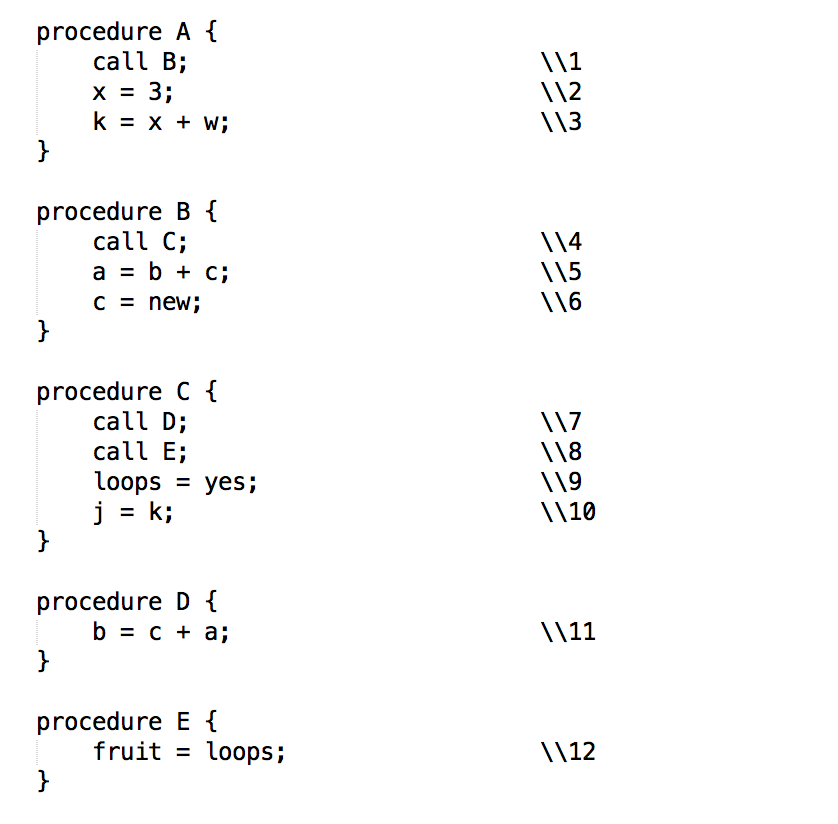
### Integration Testing

#### Parser, PKB & Design Extractor

Test Purpose: Test the interaction of Paser and Design Extractor with PKB

Required Test Inputs: PKB, Parser and Design Extractor, and a simple source code.





Expected Test Results: “OK tests” if it passes all CPPUnit asserts

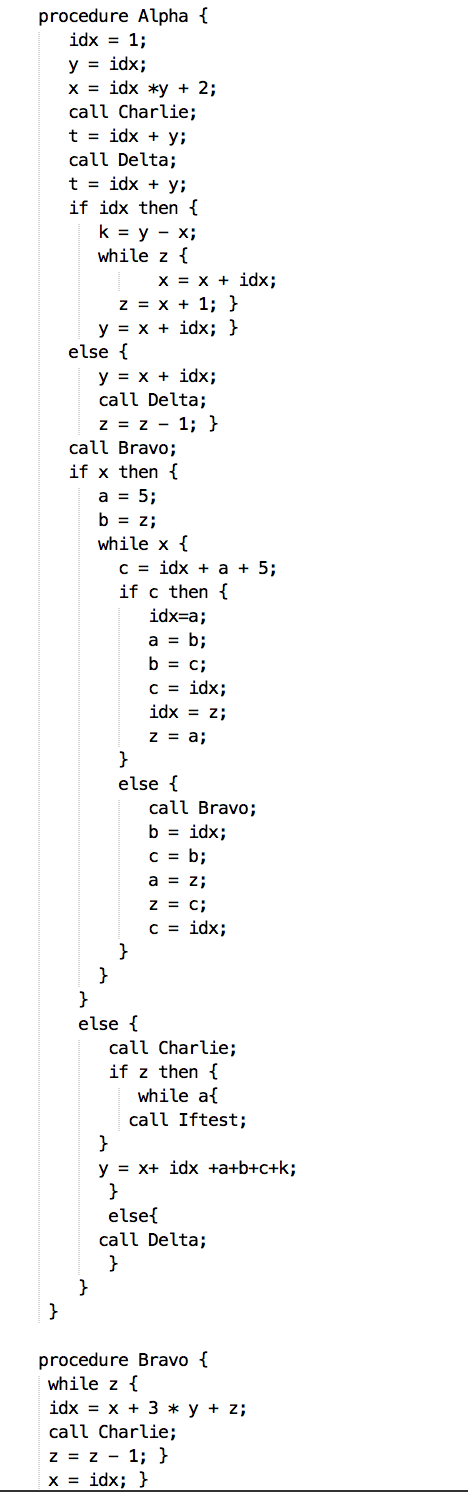
Other Requirements: None

### Validation Testing

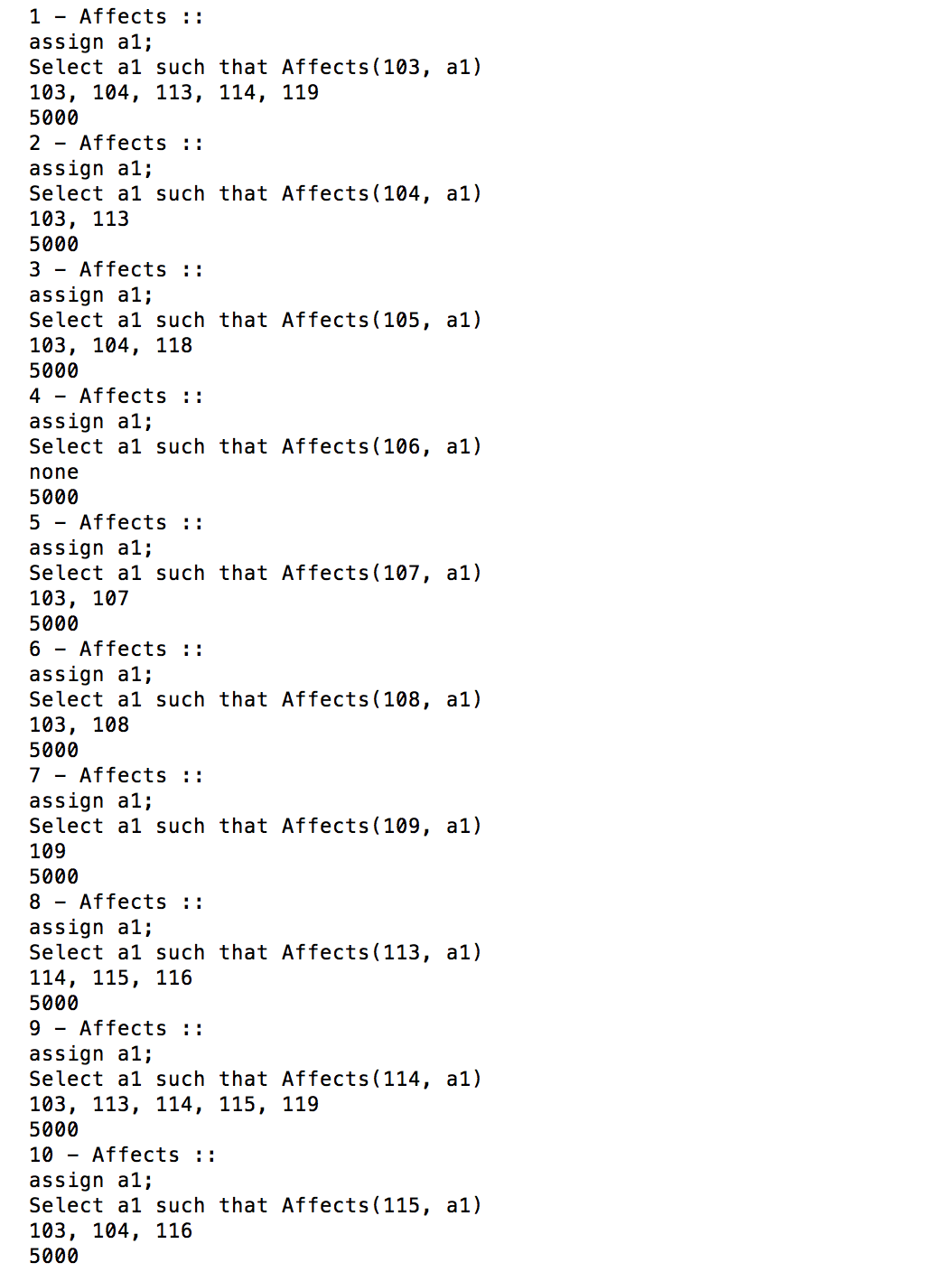
Sample 1

Test Purpose: Test Affects and AffectsStar Relationship

Required Test Inputs: The whole system. Given input is source code and query file.



Expected Test Results: As shown in figure below.

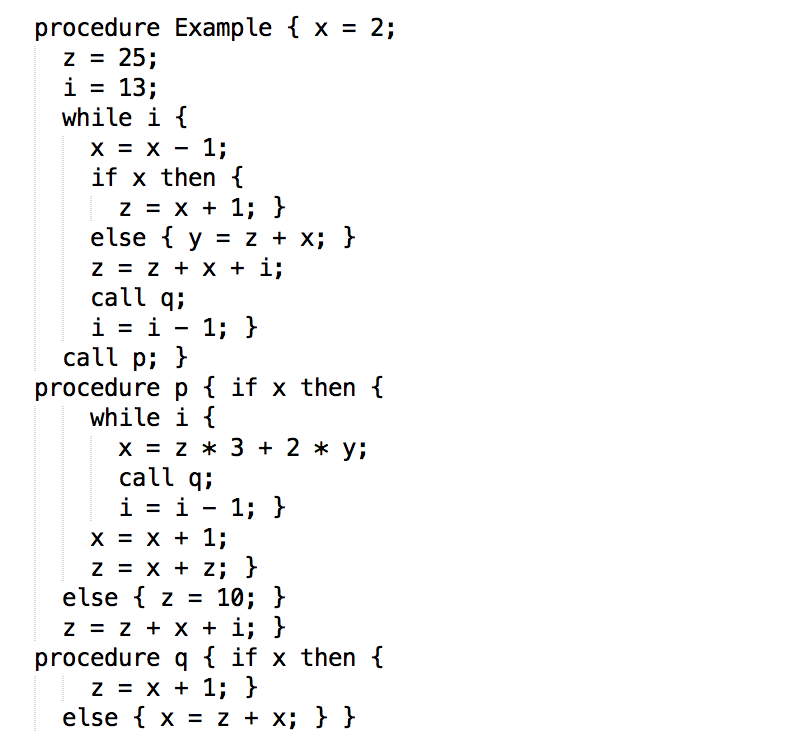


Other Requirements: None

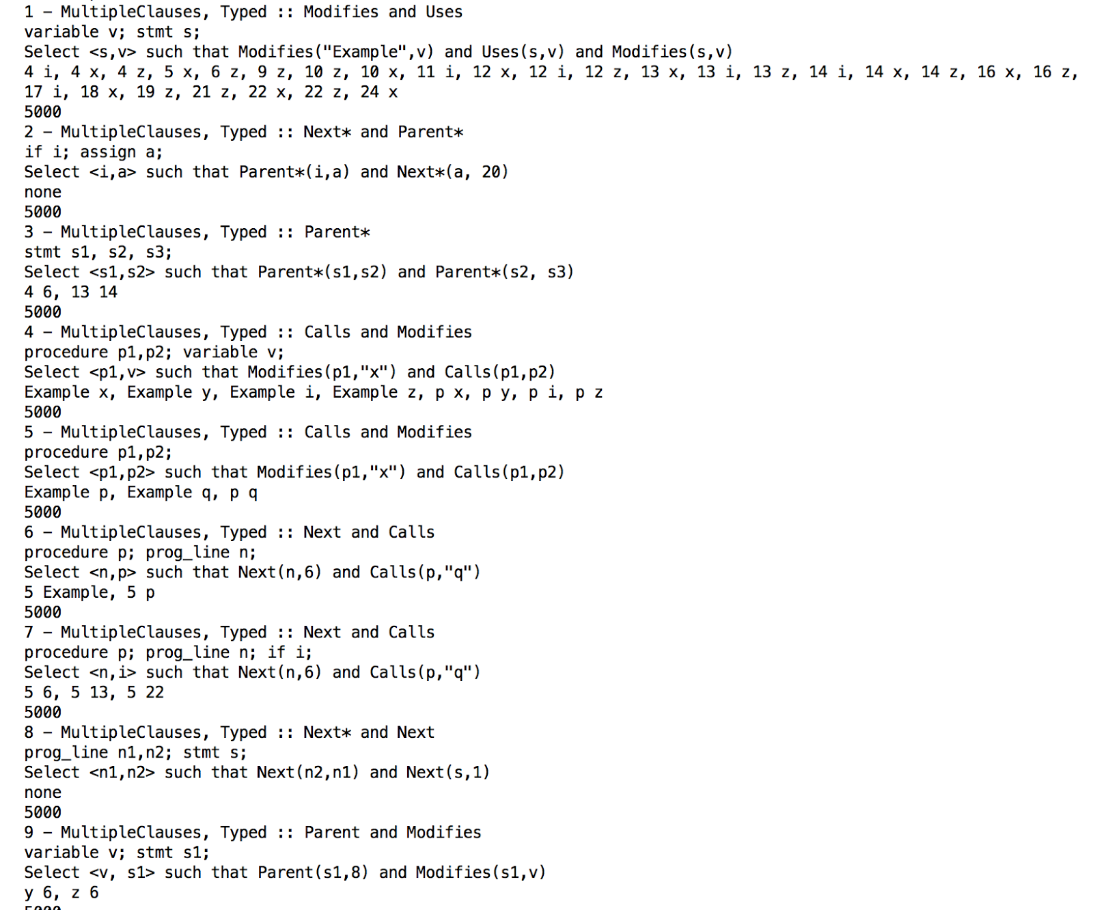
Sample 2

Test Purpose: Test returning Tuple Results

Required Test Inputs: The whole system, given input is source code and query file.



Expected Test Results: As shown in in figure below.



Other Requirements: None

# Discussion

In this iteration, the biggest challenge that we faced was implementing the Next\* and Affects/Affects\* relation correctly. Since we had not implemented the DE in iteration 2, we did not cover all of the possible cases covered for Next\* and thus for Affects as well. We realized that we missed out a major chunk of relationships between Modifies and Uses for procedure calls and Modifies/ Uses statements. Our tutor helped us pin-point this issues during our consultations. Since we added the DE, we have managed to implement the Next and Affects relationships correctly and cover all of the cases.