A means of establishing traceability based on a UML model in business application development

1st Sep. 2011
Fujitsu Laboratories, LTD.
Kyoko Ohashi, Hidetoshi Kurihara,
Yuka Tanaka, Rieko Yamamoto
Outline

- Background
- Approach for consistency checking
- Verify the approach and Add Strategies
- Consistency Checking Tool
- Evaluation and Consideration
- Conclusions and Future works
Background

Purpose: Quality improvement of business applications at an early phases of software development
- Requirements definition phase and Basic Design phase

Our company’s problems
- Member for requirements definition are different from members for basic design.
- Inconsistencies between requirements and design often occur

“Consistency checking” improves both quality of requirements and quality of design

We tried two consistency checking techniques
- Completeness verification
  - For checking shortage of design and excessive design
- Impact analysis
  - For getting design documents influenced by requirements changed
Our Goals

- To determine artifact particle granularity suitable for traceability
- To extract all candidate traceability links in order to achieve a recall rate 100%

Model-based Approach

- Define Requirements Model and Design Model
  - These models express important information of requirements/design documents
- Define traceability links between those model elements

![Diagram showing relationships between models and elements.]

Approach for Consistency Checking

- Define Requirements Model and Design Model
  - These models express important information of requirements/design documents
- Define traceability links between those model elements
Completeness Verification:

- The end of design phase, most of developer should verify completeness.
- In completeness verification, developer finds out instances with no link.

Instances of “SystemFunction” (In Requirements Model):
- Search function of vacancy:
- Find property information:
  - Registration function of progress:
  - Registration function of customer:
  - Create contract:

Instances of “Function” (In Design Model):
- General Search:
  - Manage property:
  - Show contracts:
  - General show list:
  - Manage customer:
  - Create contract data:

"Registration function of progress" does not exist in design documents.
Shortage of design?

"Calculate tax …" exists in the design documents only.
It is an excessive design. Who adds the function?
Usage of Traceability Link(2)

Impact Analysis

When requirement change occurs, developer should conduct impact analysis.

Instances of "SystemFunction" (In Requirements Model)

Oracle function of vacancy:

Find property information:

Instances of "Function" (In Design Model)

General Search:

Manage property:

Show contracts:

General show list:

Create contract:

Create contract data:

Manage customer:

If I change "Find property...", Which design instances are influenced?

By the change, 3 instances are influenced. We should check and modify those instances.
In real projects, there are thousands of requirements or design instances are defined. The number of traceability links is much more than the number of requirements or design instances.

We executed an experiment that involved setting up links manually in a large-scale project.

- Two developers with similar skills set up the links manually.
- They measured two kinds of time:
  - The time to search for a correct target of the link
  - The time to physically establish the links

Results of this experiment:

1. Both of the developers took significant time for searching.
2. The developers claimed the difficulty of setting up all links at once.

We add two strategies.
Strategy A. Limiting the Number of Candidates of Link

From the experimental result(1), the developers took significant time for searching. The ratio of the searching to the setup was roughly three to one.

So, we introduce followings:

- **Step1:** We categorize candidates by direct/indirect relation between “BusinessProcess” class and other classes in Requirements Model.
  - Only those instances that have direct/indirect relation to some BusinessProcess instances will remain as candidates.

- **Step2:** Candidates of link are ranked according to their similarity with the source artifact.
  - We use Information Retrieval method ‘Character-Based Best Match Retrieval Method’ for ranking.
Strategy B. Timing of Creating Traceability Link

- From the experimental results(1)(2), we identified two factors relevant to the heavy workload of establishing traceability links
  - **Timing**: If a developer sets up a link in a timely separated activity from the creation of the design artifact, developer has to consider a relationship with the requirement definition twice
  - **Person**: It is difficult for one developer to set up all traceability links at the end of the design phase

- So, we introduce followings:
  - **Timing**: The traceability links are set during the design phase instead of at the end
  - **Person**: Each developer in charge of a set of artifact particles establishes the traceability links for his/her particles.
We developed Consistency Checking Tool

- In our company, developers use MS Excel in early phases. Then it is difficult to apply CASE tools
- We implemented this tool with model-based approach without repository

Main functions are ‘Completeness Verification’ and ‘Impact Analysis’. We add the ‘Link Setup’ to reduce heavy workload based on the strategies A and B
Usage of ‘Link Setup’ in a Design Workflow

Customer (Requirement definition)

Supply documents

Requirement definition documents

Developer (Basic design)

Receive documents

A. Read Requirement Definition Documents

B. Create Basic Design Documents With candidates of link

C. Basic Design and set up links

D. Read Basic Design Documents

Verify Completeness of Design Documents /Impact Analysis

Requirements model

Basic design documents with candidates of link

Basic design documents with links

Design Model

Copyright 2011 Fujitsu laboratories, LTD
Usage of ‘Link Setup’ in a Design Workflow

Customer (Requirements definition)

- Supply documents
- Requirement definition documents

Developer (Basic design)

1. Receive documents
2. A. Read Requirement Definition Documents
3. B. Create Basic Design Documents With candidates of link
4. C. Basic Design and set up links
5. D. Read Basic Design Documents
6. Verify Completeness of Design Documents /Impact Analysis

Requirements model

- Basic design documents with candidates of link
- Basic design documents with links

Design Model
A. Reader for Requirements Definition Artifacts

This function creates instances based on the Requirements Model from requirements definition artifacts

【An example of a requirements definition artifact(part of a system function list)】

<table>
<thead>
<tr>
<th>ID</th>
<th>SystemFunction</th>
<th>Name of subsystem</th>
<th>Online/Batch</th>
<th>Screen/Slip/Interface</th>
<th>Type (Screen/Slip/IP)</th>
<th>ID</th>
<th>Name</th>
<th>Rank of scale and difficulty</th>
<th>Business process ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF-001</td>
<td>Search function of vacancy</td>
<td>Sale</td>
<td>Online</td>
<td>Screen</td>
<td>BP-004</td>
<td>Property</td>
<td>0</td>
<td>BP1111,BP1113</td>
<td></td>
</tr>
<tr>
<td>SF-003</td>
<td>Registration function of customer</td>
<td>Sale</td>
<td>Online</td>
<td>Screen</td>
<td>EN-002</td>
<td>Searching vacancy</td>
<td>A</td>
<td>BP1112,BP1113</td>
<td></td>
</tr>
<tr>
<td>SF-004</td>
<td>Registration function of progress</td>
<td>Sale</td>
<td>Online</td>
<td>Screen</td>
<td>EN-003</td>
<td>Registration customer</td>
<td>B</td>
<td>BP1111,BP1112,BP1113</td>
<td></td>
</tr>
</tbody>
</table>

- Create SystemFunction instances
- Set links between SystemFunction and BusinessProcess instances

【An example of Requirements Model instances(part of)】
Usage of ‘Link Setup’ in a Design Workflow

Customer (Requirements definition)

- Supply documents
- Requirement definition documents

Developer (Basic design)

- Receive documents
- A. Read Requirement Definition Documents
- B. Create Basic Design Documents With candidates of link
- C. Basic Design and set up links
- D. Read Basic Design Documents

Verify Completeness of Design Documents / Impact Analysis

Requirements model
- Basic design documents with candidates of link
- Basic design documents with links
- Design Model
B. Create Design Artifacts with Candidate Links

This function generates design artifacts from the instances of Requirements Model. And it embeds a set of instances based on Requirements Model when it creates design artifacts.

- Creates Basic Design Artifacts (MS Excel Documents).
- They embeds Requirements Model instances.

### An example of a design artifact (part of a specification of function)

<table>
<thead>
<tr>
<th>Specification of Function</th>
<th>UI01</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function name</td>
<td></td>
</tr>
<tr>
<td>Requirement_System function</td>
<td></td>
</tr>
</tbody>
</table>

This document is only template.
Usage of ‘Link Setup’ in a Design Workflow

Customer (Requirements definition)

Supply documents

Requirement definition documents

Developer (Basic design)

Receive documents

A. Read Requirement Definition Documents

B. Create Basic Design Documents with candidates of link

C. Basic Design and set up links

D. Read Basic Design Documents

Verify Completeness of Design Documents /Impact Analysis

Requirements model

Basic design documents with candidates of link

Basic design documents with links

Design Model

Copyright 2011 Fujitsu laboratories, LTD
C. Setting Up Traceability Links

This tool function helps set correct links in the previously generated design artifacts.

【The example of a design artifact(part of a specification of function)】

This document is used to define a function in design phase.

Field “Requirement System function”: For traceability links.

Show candidates of link and set selected candidates by developer.

Requirements Model instances are expressed by this tree.

Selected candidates are set this field.
Usage of ‘Link Setup’ in a Design Workflow

<table>
<thead>
<tr>
<th>Customer (Requirements definition)</th>
<th>Developer (Basic design)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Supply documents</td>
<td>Receive documents</td>
</tr>
<tr>
<td>Requirement definition documents</td>
<td>A. Read Requirement Definition Documents</td>
</tr>
<tr>
<td></td>
<td>B. Create Basic Design Documents With candidates of link</td>
</tr>
<tr>
<td></td>
<td>C. Basic Design and set up links</td>
</tr>
<tr>
<td></td>
<td>D. Read Basic Design Documents</td>
</tr>
<tr>
<td></td>
<td>Verify Completeness of Design Documents /Impact Analysis</td>
</tr>
</tbody>
</table>

Copyright 2011 Fujitsu laboratories, LTD
D. Reader for Design artifacts

- This tool function creates some instances of design instances from the design artifacts.

【The example of a design artifact (part of a specification of function)】

<table>
<thead>
<tr>
<th>Specification of Function</th>
<th>UI01</th>
<th>Real estate prop</th>
</tr>
</thead>
<tbody>
<tr>
<td>Function name</td>
<td>Customer management</td>
<td>Function ID</td>
</tr>
<tr>
<td>Requirement, System function</td>
<td>Search function of vacancy</td>
<td></td>
</tr>
</tbody>
</table>

1. 構想概要定義
1-1. 職務の概要

顧客情報に対して各種の処理を行う。
当機能は販売、仕入れ、企画、経理、の4業務について統合機能とする。
① 各業務で生成された顧客情報を作成する。
② 物件や契約をもとに検索、停止、削除を可能とする。

- Create Function instances
- Set links between Function and SystemFunction instances
Evaluation about workload reduction

We conducted a formal evaluation in order to determine, whether our approach and strategies led to a reduction in effort, while maintaining 100% recall of traceability links.

Our study used requirements and design artifacts from a real software development project.

The project size (requirements definition artifacts)

<table>
<thead>
<tr>
<th>Class name of Requirements Model</th>
<th>The number of instances (=candidates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BusinessProcess (most detail)</td>
<td>72</td>
</tr>
<tr>
<td>SystemFunction</td>
<td>397</td>
</tr>
<tr>
<td>Screen</td>
<td>526</td>
</tr>
<tr>
<td>Slip</td>
<td>145</td>
</tr>
<tr>
<td>External Interface</td>
<td>116</td>
</tr>
<tr>
<td>Conceptual Data</td>
<td>53</td>
</tr>
<tr>
<td>Total</td>
<td>1209</td>
</tr>
</tbody>
</table>

We tried to establish links between SystemFunction class instances and Function class instances.
Measurement Method

We focus the number of candidates and the time to create links for evaluating the reduction in effort.

1) Effect of models (the number of candidates)
   - Case I: The models are not used.
     We take the sum of all the instances
   - Case II: The models are used.
     We take an average of the number of instances of classes

2) Effect of categories (the number of candidates)
   - Case I: categories unused.
     The number of candidates is the number of instances of a class
   - Case II: categories used.
     The average number of instances of a class included in a business process of the most detailed

3) Effect of categories (setting time)
   - We evaluate the effect of our strategies by measuring time to set up links.
     The setting is supported by our tool
Results

1) Effect of models (the number of candidates)

<table>
<thead>
<tr>
<th>models</th>
<th>Case I</th>
<th>Case II</th>
</tr>
</thead>
<tbody>
<tr>
<td>The number of candidates</td>
<td>1209</td>
<td>201.5</td>
</tr>
</tbody>
</table>

The number of candidates was reduced to 16.7%.

2) Effect of categories (the number of candidates) & 3) Effect of categories (setting time)

<table>
<thead>
<tr>
<th>Categories</th>
<th>Case I (Unused)</th>
<th>Case II (Used)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B-1. The number of candidates</td>
<td>397</td>
<td>8.1</td>
</tr>
<tr>
<td>B-2. Setting time (sec/link)</td>
<td>23</td>
<td>15</td>
</tr>
</tbody>
</table>

The number of candidates was reduced to 2%
Setting time was reduced to about 65%
Consideration

- The number of candidates and setting time were reduced by our approach and strategies.

- Applying our approach resulted in a small enough number of candidate links, and reduced time to be effective in real software development projects.

- In the scope of SystemFunction class, Function class, and the instances related to them, recall rate was 100%.
Conclusions and Future work

- We have reported on how to establish traceability between the requirements definition and the design in a real project.

- We define traceability with models of a requirements definition and a design. The candidate links are categorized using the information of requirements definition.

- Our tool can reduce the number of candidate traceability links because it has two models and uses a link categorization strategy.
  - A developer can select the correct links easily by reviewing the limited candidate links using the models, and through categorized candidates.
  - The effect is confirmed using the requirements definition artifacts from a real software development project.

- Future works: Validate/refine our models with an additional case study, etc.
Verify the approach by recall, precision metrics

We verified traceability link by common traceability link’s metrics (recall, precision).

Metrics definition

C: set of Candidates of link.
R: set of Known links (correct links)

\[
\text{recall} = \frac{|R \cap C|}{|R|} \quad \text{precision} = \frac{|R \cap C|}{|C|}
\]

In our approach,

- Set C contains all instances of a Requirements Model class.
- Recall
  - Set R is always subset of Set C. No other classes contain R’s element.
  - So, recall rate is 100%.
- Precision
  - In real project, Set C has very large cardinality. And Set R has very small cardinality.
  - So, precision rate is very low.