TELKOMNIKA, Vol.9, No.1, April 2011, pp. 81~88
ISSN: 1693-6930
accredited by DGHE (DIKTI), Decree No: 51/Dikti/Kep/2010

Formal Specification for Spatial Information Databases Integration Framework (SIDIF)

Mustafa Man*[1], Julaily Aida Jusuh[2], Mohd Shafry Mohd Rahim[3], Mohammad Zaidi Zakaria[4]
[1]Dept. of Computer Science, Fac. of Science and Technology, Universiti Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, Malaysia, Tel: 09-6683307 fax: 609-6694660
[2]Dept. of Computer Science, Fac. of Computer Science and Information Technology, Sultan Zainal Abidin
21030 Kuala Terengganu, Terengganu, Malaysia
[3]Dept. of Computer Graphics & Multimedia, FSKSM, Universiti Teknologi Malaysia
81310 UTM Skudai, Johor Bahru, Malaysia
[4]Dept. of Fisheries Science, Fac. of Agro Technology and Food Science, Universiti Malaysia Terengganu
21030 Kuala Terengganu, Terengganu, Malaysia, Tel: 09-6683307 fax: 609-6694660
e-mail: mustafaman@umt.edu.my*[1], julaily_04@yahoo.com[2], shafry@utm.my[3], zaidi@umt.edu.my[4]

Abstrak

Makalah ini membahas validasi formal untuk Informasi Spasial Database Integrasi Framework (SIDIF). Database SIDIF adalah data persisten terorganisasi, biasanya berhubungan dengan perangkat lunak komputer yang dirancang untuk update, query, dan mengambil komponen dari data yang disimpan dalam sistem. Salah satu kesulitan yang umum dihadapi oleh pengembang adalah merancang sebuah sistem database yang kuat. Meskipun demikian, untuk menyelesaikan masalah ini, pengembang harus memfokuskan upaya mereka pada spesifikasi formal. Spesifikasi formal seharusnya mengurangi waktu pembangunan secara keseluruhan. Spesifikasi formal dapat digunakan untuk memberikan suplemen jelas dan tepat untuk deskripsi bahasa alami. Selain itu dapat divalidasi dan diverifikasi ketat menuju deteksi dini kesalahan spesifikasi. Akibatnya, untuk memvalidasi masalah ini secara formal, kita tentukan kerangka SIDIF database menggunakan bahasa Z dan membuktikan dengan menggunakan Z/EVES alat tetesan mata teorema terbukti. Dengan menggunakan alat semacam ini dapat membantu mengurangi waktu, tenaga dan kesalahan dibandingkan dengan manual membuktikan teorema yang dapat kesalahan tugas dan membosankan.

Kata kunci: *bahasa spesifikasi Z, rekayasa perangkat lunak, SIDIF, spesifikasi formal*

Abstract

This paper discusses the formal validation for spatial information databases integration framework (SIDIF). A SIDIF database is a large, organized body of persistent data, usually associated with computerized software designed to update, query, and retrieve components of the data stored within the system. One of the common difficulties faced by the developer is in designing a robust database system. Even so, in order to solve this matter, developers have to focus their efforts on the formal specifications. The formal specification is supposed to reduce the overall development time. Formal specifications can be used to provide an unambiguous and precise supplement to natural language descriptions. Besides, it can be rigorously validated and verified leading to the early detection of specification errors. Consequently, to validate this problem formally, we specify the SIDIF database framework using Z language and prove by using Z/EVES theorem proven tool. By using this kind of tools, it may help to reduce time, energy and mistake compared to manual theorem proving which can be error task and tedious.

Keywords: *formal specification, SIDIF, software engineering, Z specification Language*

1. Introduction

Software requirement specification (SRS) is a complete description of the behavior of a system to be developed. It includes a set of use cases that describe all the interactions the users will have with the software. Use cases are also known as functional requirements. In addition to use cases, the SRS also contains non-functional (or supplementary) requirements. Non-functional requirements are requirements which impose constraints on the design or implementation (such as performance engineering requirements, quality standards, or design constraints).In a few cases, the user and system requirements may be integrated into a single description. In other cases, the user requirements are defined in an introduction to the system

Received January 8th, 2011; Revised March 8th, 2011; Accepted April 8th, 2011

requirements specification. If there are a large number of requirements, the detailed system requirements may be presented in a separate document.

Specification and design are inextricably intermingled. Architectural design is essential to structure a specification and the specification process. Formal specifications are expressed in a mathematical notation whose vocabulary, syntax and semantics are formally defined [1]. Usually both the system requirements and the system design are expressed in detail and carefully analyzed and checked before implementation begins.

The necessary information content and recommendations for an organization for software design descriptions (SDDs) are described. An SDD is a representation of a software system that is used as a medium for communicating software design information. This recommended practice is applicable to paper documents, automated databases, design description languages, or other means of description. If a formal specification of the software is developed, this usually comes after the system requirements have been specifying but before the detailed system design. There is a tight response loop between the detailed requirements specification and the formal specification as shown in Figure 1 [2]. One of the main benefits of formal specification is its ability to uncover problems and ambiguities in the system requirements.

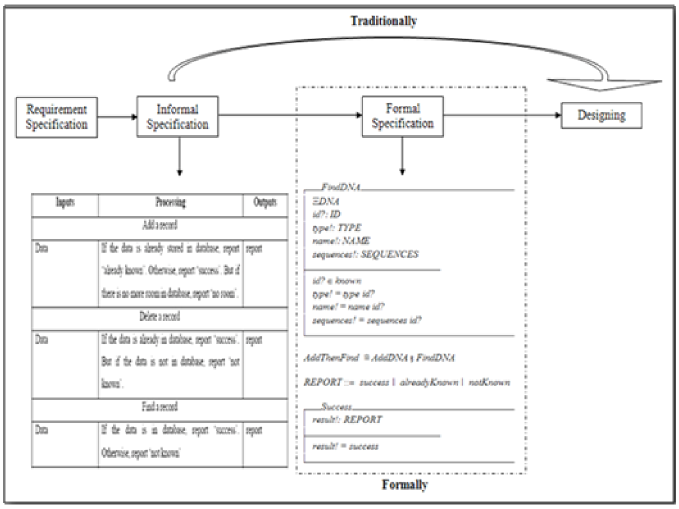

Figure 1. The Requirement Specification Process

The principal value of using formal methods in the software process is that it forces an analysis of the system requirements at an early stage. Correcting errors at this stage is cheaper than modifying a delivered system. Formal methods consist of formal specification, specification analysis and proof, transformational development and program verification. Formal specification techniques are most cost-effective in the development of critical systems where safety, reliability and security are particularly important.

Nowadays, formal proving can be done with the support of formal method tools such as theorem proven tools. Theorem proven is a tool that implements automatic theorem proving need of user support. Regularly, developers cover a long times and looping process, so there might be a great possibility of mistakes. The proofs are efficiently when it been presented in a user-friendly approach and it should not be unreasonably large. Nevertheless, a lot of the proof that involved in software validation is naturally detailed, low-level and repetitious.

So we can briefly state that it is unsuitable for human checking. Thus, formal proving supported by tool, which is not only reduce the possibility of mistakes but also not totally removes it [3]. Hence, the use of support tool is a main factor that can affect the acceptance of formal method practically [4].

The Z specification language is a way of decomposing a specification into small pieces called schemas [5]. Each piece can be linked with a commentary that gives explanation informally the importance of the formal mathematics. A schema is essentially the formal specification analogous to programming language subroutines that are used to structure a system, where the schemas are used to structure a formal specification. The Z is physically powerful on sets and functions. Generally, Z notation is used for sequential situation. We interested in using Z notation because it is a mature technique for model-based specification [6].

2. SIDIF Framework

Spatial information databases integration framework or SIDIF is a design database integration space proposed to evaluate development effectiveness artificial reef. This method is was one method regarded as a new idea to ensure effectiveness of artificial reef development project can be valued it effectiveness.

As evaluation process previously require fairly high cost required one special scuba unit to assess artificial reef levels of development and development by diving method. SIDIF was one method for information combination of two or more database which possess different scheme and also same. In this study, research only made integration of two database only namely artificial reef database (ARPOS) and fish landing database (WiFISH) [7], [8]. Before integration process both this database is conducted, various issues should be identified over proceed. A few process or move need to be taken broken up into 4 levels as follows [9]:

i. *Pre-integration*: Process to assess database environment are used example like Oracle, MySql, MSSQL, MS ACESSS and form other database.
ii. *Scheme comparison*: Scheme comparison or structure for each this database is needed to facilitate integration process conducted. In early, this process is made manually.
iii. *Intermediary software development (middleware)*: an application shall be developed for integration process the data base workable.
iv. *Post-integration*: Integration process assessment was being conducted from credibility process aspect and "*interoperability*".

The evaluation results of the two databases are found that the location (position) of artificial reef can be equal with the location of fish catches conducted. Due to this, the location based technique is a core and a fundamental to determine the effectiveness development level of the artificial reef project development at a certain location and also within the timeline. An equivalent assumption can be formed as follows:

> *Catch location (CL) is equivalent with artificial reefs development (AR). Assessment factor is dependent on catch yield number (CT) for each type of fish (FT) and comparison with artificial reef type (RT) which included. A formula can be set up here. CL ≡ AR where if found CT ≡ FT is high then AR is effective which depending on catch date (DT) made. Vice versa if CT low, then AR to be ineffective.*

Based on that assumption formula, a newly developed algorithm called location based technique can assess the effectiveness level of artificial reef development project. The integration of location based process and scheme comparison would be made by intermediary software (middleware) which shall be developed as shown in Figure 2.

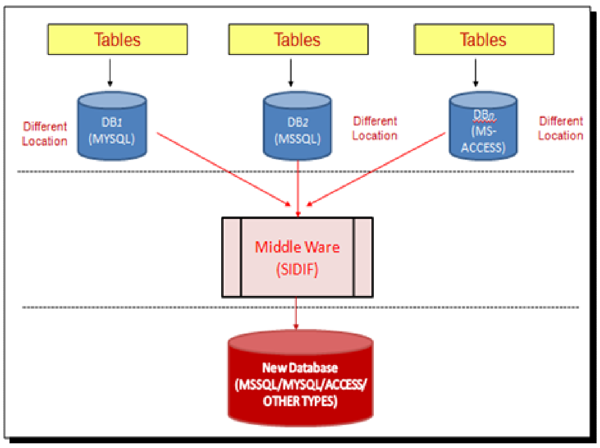

Figure 2. SIDIF architecture

Since those data in the three different databases are at different location and in different formats, then intermediary software (middleware) is very much needed for integration purposes [10]. Table 1 until Table 5 below shows the information about the data that convert to specific database type in different environment [11].

Table 1. Artificial reefs (AR) distribution information

Code	Areas	Latitude	Longitude	Type	Total	Year Built	Total Cost	Granted By
TTS01	Pulau Bidong	1000	123555	KUBOID	950	1998	49970	KWP
TTS01	Pulau Bidong	1000	123555	KUBOID	950	1998	49970	KWP
TTS16	Pulau Perhentian	200	205	SERAMIK	180	2003	343300	DANA KHAS
TTS04	Pulau Kapas	95	100	KUBOID	1400	1999	122900	KWP
..	..	..	..	..	..	..	..	..

Table 2. Daily Fish landing Data Collection

Boat No.	Landing Date (2009)	Landing Time	Fish Code	Fish Type	Total Landing	Price	Area of Fishing	Landing Port
TAP001	Dec 12	11.00	AA1	Kerisi	100	2	P.Bidong	P. Kambing
TAP001	Dec 12	11.00	AA2	Kembong	200	7	P.Kapas	P. Kambing
TAP002	Dec 13	10.00	AA1	Kerisi	100	2	P.Perhentian	P. Kambing
..	..	..	..	..	..	..	..	..

Table 3. Vessel Profile

Fisherman Name	Reg. No	Boat No	Loading	Fishing Zone	Main Fishing Equip-ment	Other Fishing Equip-ment	License no.	License expiry Date	Crew No.
Ahmad bin Ali	1112	TAP001	100	Zon A	Pukat Hanyut	Bubu	1234	20/2/2012	8
Ahmad Durah B. Sameon	2222	Tap002	500	Zon B	Pukat Tunda	Tiada	5678	23/3/2013	4
..	..	..	..	..	..	..			..

Table 4. Integration two Tables with New Information (SIDIF)

Landing date (2009)	Fish Type	Total Landing	Area of Fishing	Artificial Reef Type
Dec 12	Kerisi	100	Pulau Bidong	Kuboid
Dec 12	Kembong	200	Pulau Kapas	Kuboid
Dec 13	Kerisi	100	Pulau Perhentian	Seramik
..	…	…	…	…

Table 5. Integration of three Tables with New Information (SIDIF)

Landing date (2009)	Boat No.	Fish Type	Total Landing	Area of Fishing	Artificial Reef Type	Fishing Zone	Main Fishing Equip-ment	Other Fishing Equip-ment	Crew No.
Dec 12	TAP001	Kerisi	100	Pulau Bidong	Kuboid	Zon A	Pukat Hanyut	Bubu	8
Dec 12	TAP001	Kembong	200	Pulau Kapas	Kuboid	Zon A	Pukat Hanyut	Bubu	8
Dec 13	TAP002	Kerisi	100	Pulau Perhen-tian	Seramik	Zon B	Pukat Tunda	Tiada	4
..	..	…	…	…	…	…	…	…	…

3. Formal Specification Development using Z-eves Tools

Based on the information in Table 1 until Table 4, the formal specification is generated using Z-eves tools for formalizing the software specification for all the information. Figure 3 until Figure 5 illustrated on the schema for all the information with different type of databases in single server.

```
__KEDUDUKANTUKUN________________
kodTukun: KODTUKUN
kawasanTangkapan: KwsnTangkapan
kedudukanLatitude: KddknLatitude
kedudukanLongitude: KddknLongitude
jenisTukun: JnsTukun
bilanganTukun: BilTukun
tahunDibina: ThnBina
jumlahPeruntukan: JumPeruntukan
PembiayaanProjek: Biaya
________________________________
```

```
[KODTUKUN, KwsnTangkapan, KddknLatitude, KddknLongitude, JnsTukun, BilTukun,
 ThnBina, JumPeruntukan, Biaya]

__KEDUDUKANTUKUN________________
kodTukun: KODTUKUN
kawasanTangkapan: KwsnTangkapan
kedudukanLatitude: KddknLatitude
kedudukanLongitude: KddknLongitude
jenisTukun: JnsTukun
bilanganTukun: BilTukun
tahunDibina: ThnBina
jumlahPeruntukan: JumPeruntukan
PembiayaanProjek: Biaya
________________________________

__KedudukantukunExt______________
Kedudukantukun: 𝔽 KEDUDUKANTUKUN
________________________________

__ChangeKedudukantukun__________________________
ΔKedudukantukunExt
ΔKEDUDUKANTUKUN
x?: KEDUDUKANTUKUN
________________________
x? ∈ Kedudukantukun
θ KEDUDUKANTUKUN = x?
Kedudukantukun' = Kedudukantukun \ {x?} ∪ {θ KEDUDUKANTUKUN'}
________________________________________________
```

Figure 3. One of Z Schema for Artificial Reefs Distribution Database

```
PENDARATANIKAN
noBot: NoBot
jenisPerkakasan: JnsPkakasan
tarikhPendaratan: TarikhDarat
masaPendaratan: MasaDarat
kodIkan: KodIkan
jenisIkan: JnsIkan
jumlahHasilTangkapan: JumHslTgkapan
kawasanTangkapan: KwsnTgkapan
pusatPendaratan: PstDarat
```

```
[NoBot, JnsPkakasan, TarikhDarat, MasaDarat, KodIkan, JnsIkan, JumHslTgkapan,
 KwsnTgkapan, PstDarat]

PENDARATANIKAN
noBot: NoBot
jenisPerkakasan: JnsPkakasan
tarikhPendaratan: TarikhDarat
masaPendaratan: MasaDarat
kodIkan: KodIkan
jenisIkan: JnsIkan
jumlahHasilTangkapan: JumHslTgkapan
kawasanTangkapan: KwsnTgkapan
pusatPendaratan: PstDarat

PendaratanikanExt
Pendaratanikan: 𝔽 PENDARATANIKAN

ChangePendaratanikan
ΔPendaratanikanExt
ΔPENDARATANIKAN
x?: PENDARATANIKAN
-----------------
x? ∈ Pendaratanikan
θ PENDARATANIKAN = x?
Pendaratanikan' = Pendaratanikan \ {x?} ∪ {θ PENDARATANIKAN'}
```

Figure 4. One of Z Schema for Daily Fish Landing Database

Figure 5. Z Schema for integration of artificial reefs distribution data and daily fish landing data (two databases)

4. Result/ Output Validation

Once the formal specification and validation was successful, then coding process is done using .net programming language for developing a middleware application. Below is one module of coding for integrating two databases which is included in SIDIF framework. The output for this module is shown in Figure 6 as below.

```
//combining the data from two different databases

DataTable combine = new DataTable();
    combine.Columns.Add("TARIKHTANGKAPAN", typeof(string ));
    combine.Columns.Add("JENISIKAN", typeof(string));
    combine.Columns.Add("JUMLAHHASILTANGKAPAN", typeof(string));
    combine.Columns.Add("KAWASANTANGKAPAN", typeof(string));
    combine.Columns.Add("JENISTUKUN", typeof(string));

   for (int i = 0; i < dt1.Rows.Count;        i++)
    {
      for (int j = 0; j < dt2.Rows.Count; j++)
      {
        if (dt1.Rows[i]["TAPAKUNJAM"].ToString() == dt2.Rows[j]["KAWASANTANGKAPAN"].ToString ())
        {
          combine.Rows.Add(dt2.Rows[j]["TARIKHPENDARATAN"].ToString(),
dt2.Rows[j]["JENISIKAN"].ToString(), dt2.Rows[j]["JUMLAHHASILTANGKAPAN"].ToString(),
dt1.Rows[i]["TAPAKUNJAM"], dt1.Rows[i]["JENISTUKUN"]);
        }
      }
    }

    GridView3.DataSource= combine ;
    GridView3.DataBind();
  }
```

KOMBINASI DATA

Tarikh Tangkapan	No Bot	Jenis Ikan	Hasil tangkapan	Lokasi Tangkapan	Jenis Tukun	Zon Tangkapan	Peralatan Utama	Peralatan Tambahan	Bilangan Kru
12.12.2009	TAP001	KERISI	100	PULAU BIDONG	KUBOID	zon A	PUKAT HANYUT	BUBU	8
12.12.2009	TAP001	KEMBONG	200	PULAU KAPAS	KUBOID	zon A	PUKAT HANYUT	BUBU	8
13.12.2009	TAP002	KERISI	100	PULAU PERHENTIAN	SERAMIK	zon B	PUKAT TUNDA	TIADA	4

Figure 6. Output for databases Integration Using SIDIF

5. Conclusion

SIDIF is an architecture design which can be recycled in integration problem solutions of various databases. Location based technique for determination of the artificial reef is better to obtain satisfactory decision. Due to this, if this method is successfully deployed and implemented, then the determination of effectiveness for artificial reef development project can be valued easily without requiring high cost and can give huge impact to fishing industry in Malaysia.

This research has demonstrated that the validation on Z formal specification to the case study in fishery industry databases can reduce overall development time, can overcome ambiguity by which at the same time allow an early detection of errors (if exist). According to pass experiences, many theorems have been through a long and repetitious proving process. If the proving is done manually by humans, the possibility of mistake made is high. Using Z/EVES, not only this possibility can be reduced, the proving can be done fast and reliable. One of our future works shall deal with complete and precise specification and validation for multiple databases in distributed database systems environment.

References

[1] Sommerville I. Software Engineering, 7th ed. UK: Addison-Wesley. 2004.
[2] Mustafa M, Julaily J, Yazid MSM. Bio-Informatics: Formal Specification for DNA Databases System. LAP Lambert Academic Publishing. 2011.
[3] Bowen JP, Hinchey MG. Ten Commandments of Formal Methods. *Computer.* 1995; 28(4): 56-63.
[4] Babich F, Deotto L. Formal Methods for Specification and Analysis of Communication Protocol. *IEEE Communication Surveys & Tutorials.* 2002; 4(1): 2-15.
[5] Bowen J. Formal Specification and Documentation Using Z: A Case Study Approach. Thomson Publishing. 2003.
[6] Saaltink M. The Z/EVES System. ZUM, LNCS. 1997.
[7] Mustafa M, Yazid MSM, Khalid S, Aezwani WABW. *GIS Spatial Data Visualization Tools for Artificial Reefs Distribution.* The 3rd International Conference on Mathematics and Statistics (ICoMS-3). 2008: 25-38.
[8] Mustafa M, Yazid MSM, Aezwani WABW, Zaidi ZM. SPI: *Productive Software for Fish Landing Data Collection by Using Wireless Mobile Technology.* 1st Regional Conference on Human Resource Development (RESERD2008). Kuala Terengganu. 2008: 45-53.
[9] Parent C, Spaccapietra S. Issues and Approaches of Database Integration. *CACM.* 1998; 41(5): 166-178.
[10] Parent C, Spaccapietra S, Dupont Y. Model independent assertions for integration of heterogeneous schemas. Very Large Database Journal. 1992; 1: 81-126.
[11] Annual Fisheries 2009. Department of Fisheries, Ministry of Agriculture and Agro-based Industries. http://www.dof.gov.my/641