



Procedural Database Normalization Level Validation for Electronic Certificate Issuance

Mohammad Hanif Farhad*^a, Mr. Hafiz Mohd Sarim^b, Professor Bahari Bin
Idrus^c

*National University of Malaysia, UKM 43650 Bangi Selangor Zaba Black H 1209 National University of
Malaysia, UKM 43650 Bangi Selangor, FTSM, Black H National University of Malaysia, UKM 43650 Bangi
Selangor, FTSM, Black G*

^aEmail: hffarhad@yahoo.com

^bEmail: hafiz@ftsm.ukm.my

^cEmail: bahari@ftsm.ukm.my

Abstract

The Ministry of Communication and IT (MCIT) conducts computer training course in seven provinces of Afghanistan. These six months training courses taught computer packages for the employees of the Ministry and others sectors. At the end of this course, each participant receives a computer certificate called ECP (Electronic Certificate Printing). The current ECP issuance process faces severe problems due to the lack of procedures to validate database normalization levels in databases sent by training centers. Importation of data from non-verified database into the ECP Database has produced a mass of redundant data in an active system that cannot be redesigned from scratch. The primary consequence of this problem is that the officials in the administration office cannot find the relevant data about each student quickly, because the proper field and tables from the unverified database couldn't be found in order to find important information such as the details of the courses and marks related to a participant. The process of issuing a certificate from the ECP database results in loss of some part of information about each participant and the courses that they take, due to inconsistent database formats in the source database. As a result, the process of receiving the certificates is time-consuming, loss of money and dissatisfaction of the people who are interested to participate in MCIT courses. This research aims to improve the efficiency of ECP issuance.

* Corresponding author.

E-mail address hffarhad@yahoo.com.

The objective of the research is to develop a procedure for validating database normalization levels. This validation procedure will aid Ministry officials in the process of the recognition and the replacement of the important fields in the source database, in order to enhance the efficiency of database normalization in ECP database.

Keywords: database normalization levels; fields recognition and the replacement; database efficiency enhancement for normalization; validation procedure.

1. Introduction

Normalization is process for restructuring relation tables during database design; this process can organize data to minimize redundancy. The relationship model is at the heart of the concept of normalization and normal forms. Normalization is a procedural way to design a database which plays an important role in the maintenance of data [8]. Normalization anticipates any problems that can happen, such as redundant data and inconsistent dependencies. The method of normalization is more complex [5]. And also the process of normalization commonly makes simpler relations and reduces the danger of anomalies that may occur. Therefore, the normalization technique improves data storage [1]. Relational Databases consists of multiple, logically connected tables, where each table visually resembles a small spreadsheet. Every single column will have the same kind of information like First Name, Last Name, city, Department. Every table will have specific attributes to indicate how it relates to other tables. There relationships between tables are an important component in the database. Functional Dependency is an approach for defining the modeling relationships between tables. It is a techier derived from mathematical theory that underpins relational database theory [7]. Normalization level of database schemas, can be performed in polynomial time by using an actuate set of functional dependencies (NF) in terms of equivalence classes of attributes, for a relational schema to be in different levels of record-for, such as 2NF, 3NF or Boyce–Codd normal form (BCNF), respectively [4]. By considering and designing database forms and tables carefully, a user can minimize duplication, reduce space, make it easy to insert information, protect the data to ensure its consistency, provide faster transactions by sending less data, and can print reports of various forms [6]. Normalization is a database process which can insert data into better formed or natural relationships, such that one fact is stored in a standard database. And also the process of normalization commonly makes simpler relations and reduces the danger of anomalies that may occur. Therefore, the normalization technique improves data storage [10]. Normal forms (NF) in relation database theory represent the rules for record design. The rules corresponding for the first to fifth normal forms are presented. The normalization process depends on an analysis of functional dependencies. The concept of functional dependency is useful in the design, and analysis of a relational database. One key aspect of functional dependencies to relational databases is to represent the relation in different normal forms [9]. It plays an important role in the prevention of problems such as an insert, update or deleting anomaly depending on whether the operation which causes the error is a row insertion, update, or deletion respectively. Most of the practical recent works on database normal forms is where only the primary key is taken into account and ignoring the rest of candidate keys [2]. It is important because it emphasizes detailed careful selection and formal model combined with a formal analysis technique [3].

2. Problem Statement

The current ECP system faces problems when data is imported from databases provide by the training center. There would be difficulty to sort by names; ID number to know which allocation of time belonged to which subject's .Data which is not normalized may include the data which is contained in one or other tables. This could be bad for security reasons, disk space usage, efficiency of database updates, and, maybe most importantly, data integrity. Data before normalization is one that has not been broken down logically into smaller and manageable tables. In the current ECP system all registration information is based on the ECP application. Unfortunately the student information in ECP is full of redundant data. The objective of this presents study is to develop a procedure for validating the normalization levels of database sources, and for the recognition and the replacement of the important fields in normalization tables in order to enhance the efficiency of tables in the ECP databases.

3. System: The Electronic Certificate Current Printing (ECP) System

As mentioned earlier, the Provincial relations department of the ICT directorate of the MCIT in Afghanistan is mandated with running and operating seven training centers in different provinces of the country. Therefore, they developed a central database using SQL server 2005 and VB.net 2005 that is called the ECP system. The database is stored in the SQL server and the interface is built in VB.net. The primary cause of problem with the ECP system is that the officials in the administration office cannot find the relevant data about each student quickly and easily. Data errors in the process of issuing a certificate create incomplete reports and the ECP system cannot print the certificate directly because of this.

4. Field Selection and analysis Tables

In this case study, the approach is to procure the primary information from different book, thesis, journals and papers about Database Normalization, biased sampling methods and an open ended questionnaire, to find and discover the weaknesses about ECP database in order to get first-hand information. Normalize of the student information table to identify significant fields and preventing normalization errors will further improve the operation of the Electronic Certificate Printing system.

5. Interview

In the electronic Certificate Printing database, ECP users and students face different problems. There are two types of difficulties in the registration forms. One is for the ECP database and the other is for the user. The open-ended questionnaire is intended to find the stats of the ECP system which is related to following viewpoints:

1. Function of the course study information system: Student components, Teacher components, Subject components, Student attendance sheet components, Course components, Balance components
2. Opinions and suggestions of the Electronic Certificate Printing system stakeholders (User, Student, Kabul and provinces) about the timely arrival of Certificates, and advantages of the system on individual and also the organization.

- Opinions and suggestions of the user and student, which come from provinces in the five Zones of Afghanistan.

The Electronic Certificate Printing organizations and respondents involved in this study are.

- Province: Teacher each province has 9 teachers. Student registration is different every semester. Subject two categories: basic and advance. Basic includes the Office Package and advance includes Networking, Management and Photoshop. Attendance sheet, use for one semester. Course duration for basic is 6 months and for advance is 8 months. 'Balance' records the unpaid balance of fees.
- User in Kabul: Registers all information of students, in the ECP database which is received from the training centers in the provinces to Kabul.

Total Respondents: Five provinces that represents the five zones of Afghanistan such as Kandahar, Harat, Balk, Nangarhar, and Badakshan and a total of 15 students, 5 Teacher and the end-user which used the ECP system.

5.1 Design of the ECP System

The ECP context Diagram is shown in figure below and also the Data flow diagram is shown in figure 1. Figure 2 also describes the Main interface which is used by the end-user in Kabul.

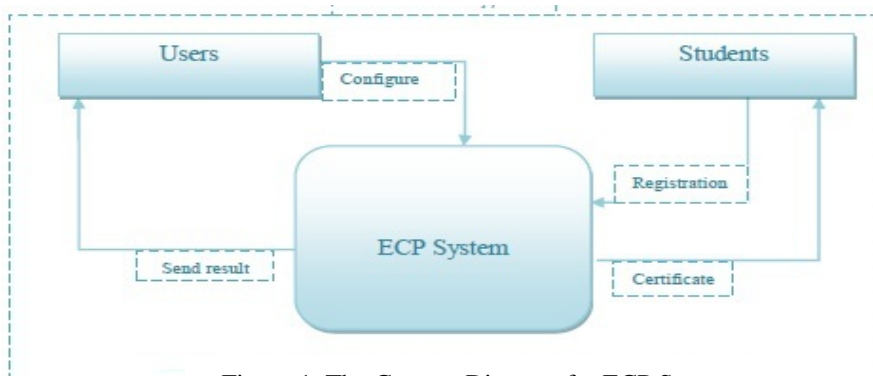


Figure 1: The Context Diagram for ECP System

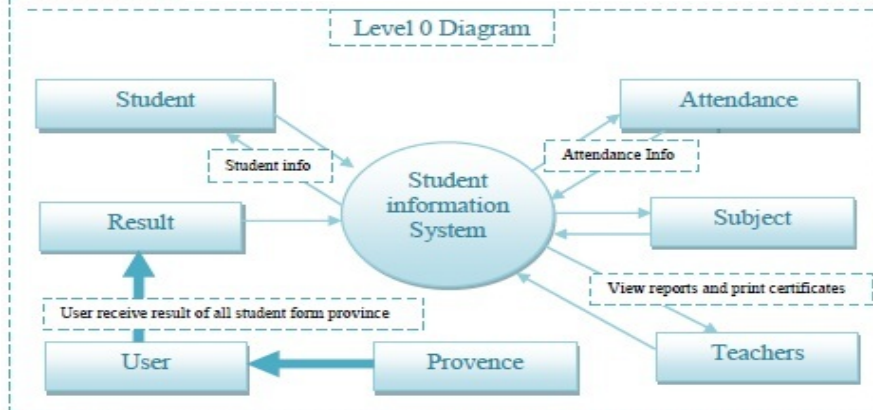


Figure 2: Sample a Result Registration Diagram at ECP system

6. Brief Summary about the ECP System

Implementing the Electronic Certificate Printing system with the database Normalization Validation procedure is first done by doing simple Open ended questionnaire, which can manage to get first-hand information about

the functions used by each province students. The verification procedure produced by this research will be used to correctly fill the ECP record and tables in the correct normal forms, even though the current system is dependent on external databases and the storage of student records is dependent on usability of the server and hardware, of the Electronic Printing System. There are weaknesses detected during registration, report, and printing of certificates. In this study a few important weaknesses are intended to be rectified through the use of the validation procedure developed in this research.

7. Relation Design and Function Dependency in the ECP System

This section describes the Relational Design of Functional Dependencies between relations graphically, using a Function Dependency diagram. The below Figure: which represent attributes and collection of attributes as rectangles, exhibit Functional Dependency using directional arrows from the contributing attributes.

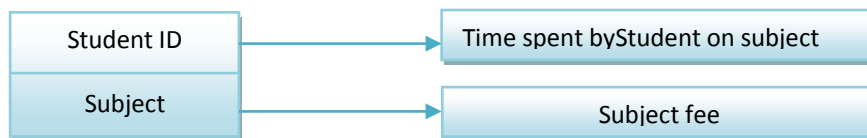


Figure3: isa sample of relation function dependency. Source: EE221 Database System PP35

The below Figure: is an example of how each of the functional dependency is depicted in the diagram. It can explain about the relationship of each type of names.

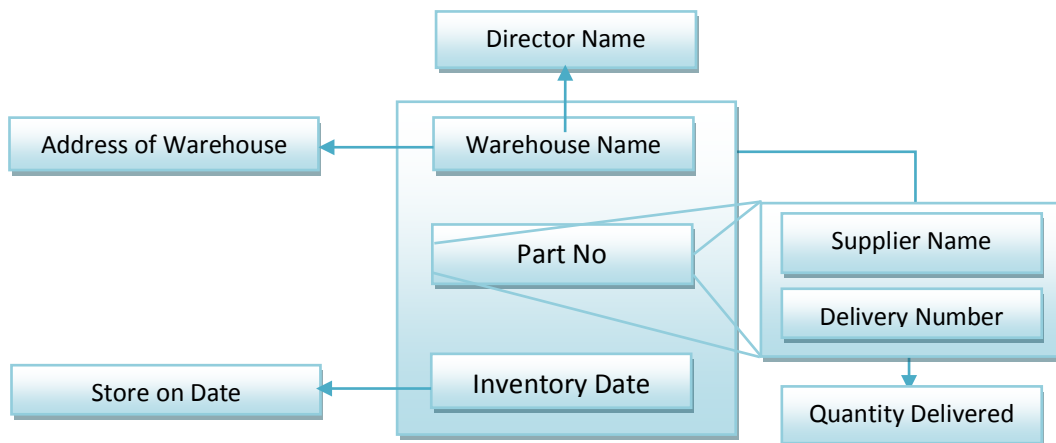


Figure4: A sample about function dependency and relation (Science 2002).Source: EE221 Database System PP

8. Recommendation for Improvement of the ECP System

Normalization-level validation procedure method operates over a complex state at a middling level of abstraction. Selecting attributes for normalization is a result of the Normalization-level validation procedure process. This chapter is dedicated to tackling the weakness of the ECP system design.

1. By introducing tables, fields, records and assumption.

2. To require normalization of the underlying ECP system.
3. To make easy to understand of information structures like (Tables, fields, columns) about the database between users and Engineers.

9. Implementation of the Normalization-Level Validation Procedure

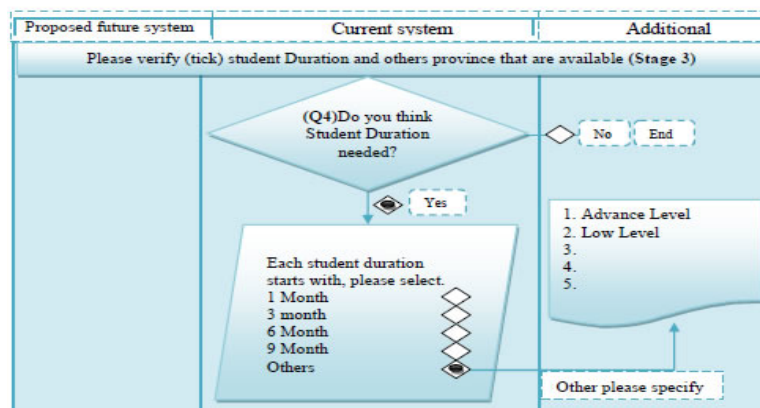
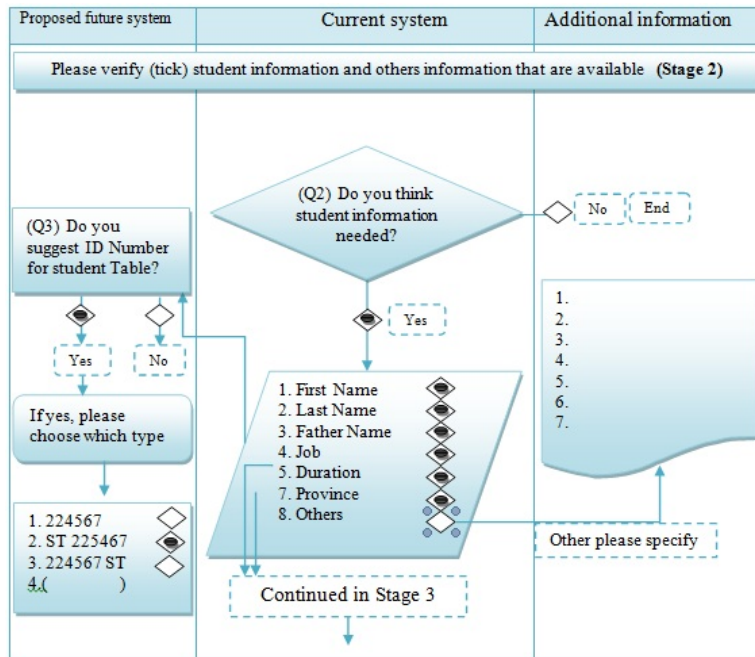
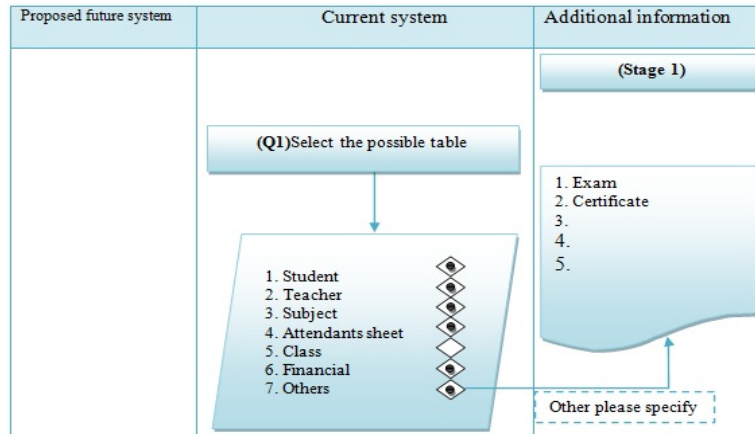
The normalization-level validation procedure uses a gradual implementation method, which is used to better recognize implementation weaknesses. The aim is that by using the normalization-level validation procedure methods then more accurate normalization can be achieved. These normalization-level validation procedure methods are a particular kind of techniques for normalization. We first look at the elementary design for an education course that is used by students in MCIT. Chapter four describes normalization-level validation procedures for their specific needs during the normalization processes. We collect the tables and fields via the normalization-level validation procedure from the ECP system, and compare their table and fields with at original design.

10. Normalize Collection and Analysis

In order to identify the number of weaknesses in the ECP system, we need to choose a system of normalized tables, fields and information that is requested in the ECP system. First, we interview some students and administrators about their studying and registration at the Provinces in Kabul. According to the Normalization-level validation Stage number 1, the following tables are collected.

11. Design and analysis normalization-level validation procedure an example a port of database such is table and student attributes

First we should develop a normalization-level validation procedure for construction database and analysis of table, field and columns of system which is different from the first one. This will only implement the existing information, database which will prepare for the student information. Accordingly only the normalization-level validation cans analysis for the construction data. Also that wills needed normalization-level validation procedure reengineering according to the designed database. It's always easy to make how a table of database will behave correct information or else. Analysis tools are not perfect, and normalization-level validation procedure is can always completely understood each selections. Therefore that, suggest for database are going to be used, you if you database designer need to do the better achieved by a very simple and easy way, that you can to ensure that your database performs well. An interview is done with a user form Ministry of Communication and Information Technology that has ECP system. In Kabul Capital of Afghanistan. But this kind of system is also in other countries, like educational department, for the education of students, etc.



Student Table, Teacher Table, Subject Table, Courses Table, Attendance Table, Financial Table, Exam Table, Certificate Table

These Normal Form table was found to be necessary in ECP system for reporting student results, so that a certificate can be printed without any problems. This also allows administrators to have access to normalized student for printing certificates on the specified time.

There are different tables in the system with specific information. Each student in the system belongs to a subject. Each subject has its own teacher to teach subject. There is an attendance sheet for student Attendance at each training center, and also student Budgets which are paid each semester. Therefore, student's registration information is collated by the center which has prepared it. Data of these centers are related to any of the student that the system is concerned with. The system records duration of training for each student information.

The following 4 stage in this case study were selected to identify the weakness of ECP system.

1. Manufacture a Process Model
2. Produce a Normalization-level validation procedure
3. Collect and Analyze for Normal forms
4. Produce a simple ER Diagram

Each stage have their own solution in order to identify weakness is in the ECP system. These stages are also identified by the Normalization-level validation procedure, for the coordination of normalization and management system. Normalization and management system are organized based on the types of table in the ECP system which has been mentioned previously. All solutions in this stage are concerned with the Normalization-level validation procedure that should be prepared for it, just in case the process model is not relevant. The normalization-level validation procedure solutions are owned by the normalization Management system.

12. Analysis of the Normalization Level Validation Procedure

This chapter has identified the weaknesses of the database and deficiency which has been obtained after the implementation of the normalization-level validation procedure, and then a question appears: what solutions can be constructed from the normalization-level validation procedure, which will handle these implementation weaknesses. There are three extreme points in the, implementation of the normalization-level validation procedure. One is the complete fit of validation procedure for the correct insertion of information in the system; the second is the identification of significant fields and records to make insertion very easy for the user. The other third suggestion is to find the relationship and extra table for each drop down list. The first extreme point has more deficiencies and weaknesses of implementation. Therefore, this three proposition is objects for this case study. One is for the complete fit of normalization-level validation for the correct question in all stages of normal forms. But the second case needs complete customization. So Electronic Certificate Printing needs a custom solution for solve the weakness. New information should be collected for all tables of at database by using the normalization-level validation procedure. This case can also achieve the objective of this case study.

13. Necessary Suggestions new Table, Fields and Columns for Registration

The first stage of the normalization-level validation procedure, discusses availability and shortage of tables, fields and records which have weakness with the normalization ECP system. Therefore, we can guarantee that all table which are obtained by normalization-level validation, will be necessary. Thus, shortage and improper repeating values violates normalization during insertion of information into the system. The main weakness is

where the inaccurate formats of data. So in the suitable table, fields can be found to prevent insertion errors, therefore the weakness of system can be solved. At present, our concerned with locating the extra fields and columns are also proposed for clarifying some columns, which is required for different tables or fields. Thus the context, customization of database for extra tables, fields and columns needs to be suggested for the designer.

14. Findings

The selected methods are interrogative techniques for finding and increasing database normalization levels using a data approach. The method aims to diagnose the possibility of updating or adding a new table, fields or others, to maintain information on a particular topic for any table's columns. The questions represent techniques to fix the data structure to obtain a particular normalization-level in the ECP database.

The Normalization-level validation procedure is a fault avoidance technique which can very easily be used to reduce errors introduced by the designer, to prepare a new system without any future problems, especially at the earlier stages of usage. It is able to remove insertion errors from the ECPs table.

15. Conclusion

The Normalization-level validation procedure plays a key role in the process of organizing the fields in a table and its dependencies. This procedure for validation of database normalization process been divided into different stages according to the type of input and relation data. Throughout the case study, we have discovered that it is not easy to create a formal procedure for validation of database normalization process which is suitable for all kinds of input and relational data. Each kind of input and relation data or information has its own normalization-validation procedure process. Therefore, there is still a lot of opportunity for research using this method.

References

- [1] A. Embong, Normalization .Database system.Kuantan.Prnerbit.2010, PP. 215-179-170-100-99-98-21.
- [2] A. Jaya. and G Sunitha. A Knowledge Based Approach for Automatic Database Normalization, International Journal of Advanced Research in Computer Engineering &Technology (IJARCET) Volume 2. No 5. May.2013,PP. 1816.
- [3] Adacore, Yannick Moy, Emmanuel Ledinot. Dassault. Hervé Delseny, 2011.
- [4] C. Vitalie. Determination of the normalization level of database schemas through equivalence classes of attributes, Computer Science, Journal 2008, Moldova, vol.17, no,2(50), PP. 125.
- [5] E.M. Arvacheh. A Study of Segmentation and Normalization for Iris Recognition System. Thesis Presented to the University of Waterloo. Ontario, Canada, 2006, PP. 5-6.
- [6] J. Morufu .SQL Database, Information Technology. SQL Database, thesis. June. 2010, PP. 16-17

- [7] J. Patel. Turbocharging DBMS Buffer Pool Using SSDs Jaeyoun do University of Wisconsin-Madison. DonghuiZhang MicrosoftJimGraySystems Lab. University ofWisconsin Madison11.June, 2010, PP 12-16.
- [8] M. Demba. Algorithm for Relation Database Normalization up to 3N International Journal of Database Management Systems Department of Computer Science & Information. Aljouf University Sakaka. Kingdom of Saudi Arabia IJDMS Vol.5. No.3. June 2013. PP 39.
- [9] M. Patwardhan. Dhabe, and Asavari. LEVEL Diagrammatic approach for Complete Automation of Relation DatabaseNormalization at Conceptual Level,International Journal. Nov 2010, PP. 141-142- 143.
- [10] P.B Alappanavar. Radhika. G and Srishti Hunjan. Automating The Normalization Process For Relational Database Model, International Journal February 2013, PP. 1826-1827.