New Reverse Engineering Tool to Transform Oracle10g Code to Entity Relationship Model

Nada N. Saleem

Saad Ahmed Th.

Nada_N_S@uomosul.edu.iq

College of Computer Sciences and Mathematics

University of Mosul, Mosul, Iraq

Received on: 11/03/2012

Accepted on: 28/06/2012

ABSTRACT

Reverse engineering is one of the important operations performed on programming codes to get their design presented graphically. Reverse engineering is performed on various types of databases to get their particular designs represented by different models of relationships.

In this paper, we found that a tool for performing reverse engineering on Oracle10g database located inside Oracle10g server, It has been designed in order to get its entity relationship model. Oracle10g data dictionary has been considered in performing reverse engineering. This dictionary includes all information that describes databases which is utilized for getting the important information relevant to the tables. It was necessary for presenting the entity relationship model. XML tags have been used in presenting the information of the data dictionary. These tags were utilized in creating the graphical entity relationship model.

Keyword: Oracle10g, Reverse engineering, Entity Relationship model, XML

نهج جديد لتحويل شفرة أوراكل 10g إلى نموذج العلاقات الكيانية

سعد أحمد نياي

ندى نعمت سليم

كلية علوم الحاسوب والرياضيات

جامعة الموصل، الموصل، العراق

تاريخ قبول البحث: 2012\06\28

تاريخ استلام البحث: 2012\03\11

المخلص

تعد الهندسة العكسية من العمليات المهمة التي يتم تنفيذها على الشفرات البرمجية للحصول على تصاميم من تلك الشفرات والتي يتم تمثيلها بشكل رسومي، يتم تنفيذ الهندسة العكسية على قواعد البيانات بأنواعها المختلفة للحصول على التصاميم الخاصة بقواعد البيانات والمتمثلة بنماذج مختلفة من العلاقات.

تم في هذا البحث تصميم وتمثيل أداة لتنفيذ الهندسة العكسية على قواعد بيانات (Oracle10g) والتي تكون متواجدة داخل مزود خدمة (Oracle10g) للحصول على أنموذج العلاقات الكيانية (Entity Relationship Model) الذي يمثلها، كما تم الاعتماد على قاموس بيانات (Oracle10g) في تنفيذ الهندسة العكسية إذ يضم قاموس البيانات جميع المعلومات التي تصف قواعد البيانات والتي يتم الاستفادة منها في الحصول على المعلومات المهمة المتعلقة بالجدول والتي من الضروري الحصول عليها لتمثيل أنموذج العلاقات الكيانية. تم الاعتماد على

وسوم لغة التوصيف الموسعة (XML) في تمثيل معلومات قاموس البيانات اذ تم الاعتماد على تلك الوسوم في تكوين انموذج العلاقات الكيانية الرسومي.

الكلمات المفتاحية: الهندسة العكسية, اوراق 10g, نموذج العلاقات الكيانية, وسوم لغة التوصيف الموسعة.

1- مقدمة

وفقاً لمعجم (IEEE) تعرف الهندسة العكسية على أنها عملية استخلاص برنامج نظام المعلومات من شفرة المصدر [1] ، وتعرف أيضاً على أنها جزء من البرمجيات التي تتكون من استعادة أو إعادة البناء الوظيفي لمواصفاتها التقنية [2,3]، تنفذ الهندسة العكسية من خلال عملية التعرف على مكونات النظام والعلاقات التي تربط تلك المكونات وتكوين تمثيل للنظام في شكل آخر [1]، كما تعد من أهم فروع التصميم الهندسي وتصنيع التطبيقات إذ تعد خطوة مهمة في دورة تطوير المنتج واستخدامها يؤدي إلى انخفاض كبير في الوقت والتكاليف للمنتج [4,5]. تعمل أدوات الهندسة العكسية على اخذ شفرة المصدر (Source Code) بوصفها ادخالاً للأداة ومن ثم توليد تحليل رسومي مهيكلي (structured Graphical analysis) ونماذج التصميم (design models) ومعلومات أخرى عن التصميم [6]. هنالك نوعان من المسوغات الأساسية لتنفيذ الهندسة العكسية وكما يأتي:

أولاً: يتم تنفيذ الهندسة العكسية من اجل فهم كيف يمكن لبرنامج الحاسوب أن يعمل فعلاً.

ثانياً: يتم تنفيذ ذلك من اجل تمكين المبرمج من فهم لماذا البرنامج لا يعمل بشكل صحيح.

في الممارسة العملية فان تطبيق الهندسة العكسية يتم من اجل مجموعة من النقاط من أهمها: [7]

1. وضع برنامج جديد يتعامل مع البرنامج الأصلي.
2. وضع برنامج جديد يتعامل مع بيانات البرنامج الأصلي.
3. إنشاء الوثائق التي تمكن المبرمجين من فهم أفضل للبرنامج الذي تمت عليه الهندسة العكسية.
4. اختبار أمانية البرنامج والتحقق من احتمال خرق الأمانة.

2- الأعمال السابقة

قام الباحث (Christian SOUTOU) في عام (1998) بتنفيذ الهندسة العكسية على قواعد البيانات العلائقية SQL92 ولكن من ابرز عيوب هذا النوع من قواعد البيانات هو عدم دعمها لتعريف المفتاح الأجنبي الذي يعد الأساس في تكوين العلاقات في أنموذج العلاقات الكيانية [8]. كما قام الباحث (Fawaz A. Masoud) في عام (2005) بتنفيذ الهندسة العكسية على قواعد بيانات Access بالاعتماد على مبدأ استخلاص البيانات من الشفرة البرمجية للغة الاستعلامات المهيكلة الخاصة بقواعد بيانات Access، وكان من ضمن توصيات الباحث العمل على تنفيذ الهندسة العكسية على قواعد بيانات Oracle [3].

وقام الباحث (Massimo Franceschet) في عام (2009) باقتراح اسلوب لتحويل أنموذج العلاقات الكيانية الى لغة التوصيف الموسعة من خلال مجموعة من العمليات والتي تؤدي الى تكوين وسوم تصف جميع مكونات انموذج العلاقات الكيانية [9].

3- لغة التوصيف الموسعة (XML) و نمذجة البيانات

تم تطوير لغة التوصيف الموسعة XML بوصفها صيغة لتمثيل البيانات لاسيما في صفحات الويب، وقيمة عملها الأساسي هو توفير معلومات حول معنى البيانات بالطريقة نفسها التي توفرها HTML مما أدى إلى

اعتمادها على نطاق واسع بوصفها صيغة لنقل البيانات بين التطبيقات والمؤسسات وتطوير مجموعة متنوعة من الأدوات لتوليد بيانات XML ومعالجة البيانات بتنسيق XML، وقد أدى نجاح XML في هذه الأدوار إلى استخدامها تنسيقاً لتخزين البيانات. ومهما كان للغة التوصيف (XML) من نقاط قوة ونقاط ضعف أخرى، فإنها لا تزال الحاجة إلى فهم صحيح لمتطلبات البيانات. [10]

4- قاعدة بيانات Oracle

يوجد العديد من أنظمة إدارة قواعد البيانات العلائقية والتي تسمح بالوصول إلى البيانات ومعالجتها ومن أبرز تلك الأنظمة هي قاعدة بيانات أوراكل 10g والتي تعد من الإصدارات الرئيسية لإدارة قواعد البيانات العلائقية التي تقلل من الوقت، التكلفة والتعقيد في إدارة قواعد البيانات من خلال إدخال بعض ميزات الإدارة الذاتية [11].

4-1 قاموس بيانات Oracle

تحتوي أنظمة أوراكل على قاموس بيانات يحتوي جميع المعلومات التي يتم استخدامها لإدارة الكائنات (Objects) الموجودة في قاعدة البيانات، على الرغم من أن قاموس البيانات يتم استخدامه في نطاق مديري قواعد البيانات (Database Administrators) لكنه أصبح الآن مصدراً مفيداً لمعلومات قيمة للمطورين والمستخدمين النهائيين. [12].

5- وصف أداة الهندسة العكسية المقترحة

تتلخص الهندسة العكسية التي يتم تنفيذها من خلال الاداة المقترحة بتحويل قواعد بيانات Oracle 10g من داخل مزود الخدمة والتي تكون في مرحلة التنفيذ إلى مرحلة التصميم المتمثلة بأنموذج العلاقات الكيانية ER Model ، إذ أن من الصعب على مصممي قواعد بيانات Oracle 10g فهم ما تحويه قاعدة بيانات أو مستخدم معين من (عدد الجداول - أسماء الجداول - عدد حقول الجداول - أنواع الحقول - أسماء الحقول - القيود المفروضة على الحقول كالمفتاح الأساسي والثانوي - أنواع العلاقات بين الجداول - درجات العلاقات) إذ أن هذه المعلومات موجودة في قاموس بيانات Oracle 10g ومن خلال محرر لغة الاستعلامات المهيكلة الخاص بـ Oracle 10g (SQL Plus) يمكن الاستعلام عن تلك المعلومات لتظهر بصورة نصوص تمثل معلومات تصف ما يحويه مستخدم معين، وفهم هذه النصوص يحتاج إلى خبرة كبيرة في قواعد بيانات Oracle 10g، ولجعل هذه النصوص مفهومة من قبل أي مصمم قواعد بيانات تم اقتراح طريقة للأداة تعمل على تحويل هذه النصوص إلى أنموذج العلاقات الكيانية الذي يمثل قاعدة البيانات في مرحلة التصميم.

5-1 تحليل عملية الهندسة العكسية في الأداة المقترحة

لتنفيذ عمليات الهندسة العكسية من خلال الأداة المقترحة يجب أن يكون مزود خدمة Oracle 10g مثبناً على الحاسوب بالإضافة إلى تثبيت الأداة المقترحة، تقوم الأداة بالاتصال بمزود خدمة Oracle 10g وبعد تحقق الاتصال يتم تحديد إحدى قواعد البيانات (مستخدم) التي تم تكوينها مسبقاً وتنفيذ جمل الاستعلام الخاصة بقاموس بيانات Oracle 10g وجلب جميع المعلومات الضرورية لتمثيل قاعدة البيانات بأسلوب أنموذج العلاقات الكيانية والمتمثلة بـ (أسماء الجداول - أسماء الحقول لكل جدول - أنواع الحقول - طول بيانات الحقول - القيود المفروضة على الحقول كالمفتاح الأساسي والثانوي - قيد التواجد لحقول المفتاح الثانوي)، يتم الاستفادة من المعلومات السابقة

في تكوين أنموذج العلاقات الكيانية وكما موضح بالجدول (1) الذي يبين معلومات قاموس البيانات وما يماثلها في أنموذج العلاقات الكيانية:

الجدول (1) معلومات قاموس البيانات وما يماثلها في أنموذج العلاقات الكيانية

معلومات قاموس البيانات	ما يماثلها في أنموذج العلاقات الكيانية
أسماء الجداول	أسماء الكيانات
أسماء الحقول	أسماء الصفات
قيود المفتاح الرئيسي والثانوي	العلاقات
قيود التواجد لحقول المفتاح الثانوي	التواجدات للعلاقات
عدد حقول المفتاح الثانوي في الجدول	درجة العلاقة

تمر عملية تحويل معلومات قاموس بيانات Oracle 10g إلى أنموذج العلاقات الكيانية بمرحلة وسطية وهي تحويل معلومات قاموس البيانات إلى وسوم لغة التوصيف الموسعة الخاصة بالأداة المقترحة وتتم هذه العملية بوضع معلومات قاموس البيانات التي يتم الاستعلام عنها بمجموعة بيانات (Data Set) ومن ثم استخدام دوال تحويل مجموعة البيانات إلى وسوم لغة التوصيف الموسعة والتي يدعمها Visual Studio 2008.NET. تشمل عملية التحويل جميع المعلومات التي تم الاستعلام عنها من قاموس البيانات ولكن سيتم تنظيمها بوسوم لغة التوصيف الموسعة مما يجعل من السهل ترتيب الوسوم الناتجة بما يتناسب مع الوسوم الخاصة بالأداة المقترحة وإضافة بعض الوسوم الضرورية والخاصة بإعدادات تكوين مساحة العمل التي سيتم إظهار وتمثيل التصميم الرسومي لأنموذج العلاقات الكيانية عليها.

2-5 خوارزمية الأداة المقترحة

تتلخص خوارزمية الهندسة العكسية للأداة المقترحة بالخطوات الآتية:

الخطوة الأولى: تحديد اسم إحدى قواعد البيانات المكونة مسبقاً داخل مزود خدمة Oracle 10g. الخطوة الثانية: فتح الاتصال بين الأداة المقترحة ومزود خدمة Oracle 10g عن طريق مكتبات الاتصال التي يدعمها V.S.2008.NET.

الخطوة الثالثة: تنفيذ استعلام قاموس البيانات الخاص بجلب المعلومات الضرورية من قاموس بيانات Oracle 10g وكالتالي:

```
SELECT TABLE_NAME , COLUMN_NAME , DATA_TYPE ,
DATA_LENGTH , NULLABLE , CONSTRAINT_NAME FROM
USER_TAB_COLUMNS, USER_CONSTRAINTS
```

الخطوة الرابعة: وضع المعلومات المستخلصة من الخطوة الثالثة في مجموعة بيانات (Data Set).

الخطوة الخامسة: تحويل مجموعة البيانات إلى وسوم لغة التوصيف الموسعة وتحويل الوسوم الناتجة إلى ما يناظرها من الوسوم الخاصة بالأداة المقترحة.

الخطوة السادسة: إضافة وسوم تهيئة مساحة العمل في الأداة المقترحة التي سيتم رسم وعرض أنموذج العلاقات الكيانية من خلالها ومن هذه الوسوم:

<SoftwareEngineeringMosul>

<Project>

<ProjectName>New Reverse Engineering Project</ProjectName>

<ProjectAuthor>Saad</ProjectAuthor>

</Project>

</SoftwareEngineeringMosul>

الخطوة السابعة: استعراض أنموذج العلاقات الكيانية بشكل رسومي من خلال الأداة المقترحة والذي يمثل قاعدة البيانات التي تمت عليها الهندسة العكسية.

مما سبق شرحه عن وصف للأداة المقترحة يمكن تلخيص الشكل العام للخوارزمية التي تم بناؤها للأداة المصممة في الشكل (1) الذي يوضح المخطط الانسيابي للأداة.

الشكل (1) المخطط الانسيابي للهندسة العكسية في الأداة المقترحة.

3-5 المحرر الرسومي للأداة المصممة

تم تصميم الأداة المقترحة ومحررها الرسومي باستخدام لغة (C#) ضمن Visual Studio 2008.NET إذ يوضح الشكل (2) المحرر الرسومي الذي تم تصميمه لرسم أنموذج العلاقات الكيانية من خلال تنفيذ الهندسة

العكسية على احدى قواعد البيانات التي تم تكوينها مسبقا داخل مزود خدمة Oracle10g والتي تمثل قاعدة البيانات التعليمية (Northwind).

الشكل (2) أنموذج العلاقات الكيانية الناتج عن الهندسة العكسية لقاعدة بيانات (Northwind) باستخدام الأداة المقترحة.

6- الاستنتاجات

- تم في هذا البحث استنتاج ما يأتي:
- تدعم الأداة المقترحة تنفيذ عمليات الهندسة العكسية على قواعد بيانات Oracle 10g وتحويلها بشكل آلي إلى أنموذج العلاقات الكيانية.
 - يعد قاموس بيانات Oracle 10g المصدر الأساس في الحصول على معلومات وصف تفاصيل أية قاعدة بيانات يتم الاستعلام عنها من خلال مزود خدمة Oracle 10g.
 - لغة التوصيفات الموسعة (XML) مهمة جدا عند بناء أي أداة تحتاج إلى التمثيل الرسومي إذ تعمل على تحويل الأوصاف الرسومية التي لا يمكن التعبير عنها برمجيا إلى أوصاف على شكل نصوص يمكن التعبير عنها برمجيا بعد تحليل تلك النصوص.

7- التوصيات

- من أهم التوصيات لتطوير عمل الأداة هي :
1. تنفيذ عمليات الهندسة العكسية لتحويل شفرات لغات أخرى مثل (sql server-My sql) إلى أنموذج العلاقات الكيانية أو أي أنموذج آخر.
 2. تنفيذ عمليات الهندسة العكسية لتحويل شفرات لغات البرمجة الكيانية الموجهة OOP إلى أنموذج لغة النمذجة الموحدة UML أو أي أنموذج آخر.

المصادر

- [1] Ian Sommerville, (2006) "Software Engineering" ,China Machine Press, China.
- [2] Abraham Silberschatz, Henry F. Korth and S. Sudarshan, (2006) "Database System Concepts", McGraw-Hill.
- [3] Fawaz A. Masoud, Heba_tallah Khattab, and Mahmoud Al-Karazoon, (2005), "University of Jordan Case Tool (Uj-Case-Tool) for Database Reverse Engineering", World Academy of Science, Engineering and Technology, Vol. 9, p.p 28-31.
- [4] Eyup Bagci, (2009), "Reverse engineering applications for recovery of broken or worn parts and re-manufacturing: Three case studies", Advances in Engineering Software(Elsevier), Vol.40 p.p 407–418.
- [5] Varady T, Martin RR, Cox J., (1997), "Reverse engineering of geometric models- an introduction", Computer Aided Design Vol. 29 (4), p.p 255–68.
- [6] Pressman, Roger S., (2001) "Software engineering: a practitioner's approach", McGraw-Hill series in computer science, Boston.
- [7] Leif Gamertsfelder, (2003), "Software reverse engineering – the current state of Australian law", Computer Law & Security Report (Elsevier),Vol. 19 no. 5 p.p 394-400.
- [8] Christian Soutou, (1998), "Inference of Aggregate Relationships through Database Reverse Engineering", 17th International Conference on Conceptual Modeling, Singapore, Springer vol. 1507 p.p. 135-149.
- [9] Massimo Franceschet, Donatella Gubiani, Angelo Montanari, Carla Piazza, (2009), "From Entity Relationship to XML Schema: a graph-theoretic approach", Proceedings of the 6th International XML Database Symposium on Database and XML Technologies, Springer-Verlag Berlin, p.p. 165-179.
- [10] Boyd L. Summers, (2011), "Software Engineering Reviews and Audits", Taylor and Francis Group, Boca Raton.
- [11] Rick Greenwald, Robert Stackowiak, Jonathan Stern, (2004), "Oracle Essentials: Oracle Database 10g, 3rd Edition", O'Reilly publisher.
- [12] Kevin Loney, (2004), "Oracle Database 10g The Complete Reference", McGraw-Hill series in computer science, Boston.