

A Genetic Based Approach for Reducing Null Values in Object-Oriented Databases

By

Sanad Abdullah Mohammad Al-Afghani

Supervisor

Dr. Said Ghoul

**This Thesis was Submitted in Partial Fulfillment of the
Requirements for the Master's Degree in Computer Science**

**Deanship of Academic Research and Graduate Studies
Philadelphia University**

June 2008

Abstract

In databases, the *Null Values* considered as one of major problems; where it may affects results accuracy and database consistency. This problem affects all databases types such as relational, object oriented, object relational. Researchers are unanimous on three reasons that may causes null values for attributes: value unknown, value unavailable or inapplicable attribute.

But they didn't observe that *the presence of alternative attributes for the same entity may cause null values*, which can be solved using the *Alternation* and the *Selection* concepts.

The *genetic approach* is one of the approaches that used the alternation and the selection concepts, it aimed to present the semantic classification in object-oriented model based on alternation concept, then interpreting collection of *genetic programs* to apply the selection concept for properties.

الطريقة الجينية للتقليل من القيم الفارغة
في قواعد البيانات الشئية

من قبل

سند عبد الله محمد الأفغاني

باشراف

د. سعيد غول

قدمت هذه الرسالة أستكمالاً لمتطلبات
الحصول على درجة الماجستير في علم الحاسوب

عمادة البحث العلمي والدراسات العليا

جامعة فيلادلفيا

حزيران 2008

ملخص

تعتبر القيم عديمة الوجود (Null Value) من أهم المشاكل التي تواجهها قواعد البيانات، فقد أجمع الباحثون على ثلاثة أسباب لهذه المشكلة: القيمة غير المعرفة، القيمة غير المتواجدة والخصيصة (Attribute) غير المتوافقة. لكنهم لم يلاحظوا أن وجود خصائص بديلة (Alternative Attribute) لنفس الكينونة (Entity) تسبب هذه المشكلة أيضا، حيث أن هذه البدائل قد تكون موجودة في نفس المجموعة (Class) أو تكتسب جراء عملية الوراثة المتعددة (Multiple Inheritance). حيث يمكن حل هذه الحالة باستخدام مفهوم البديل (Alternation) والاختيار (Selection).

من إحدى المنهجيات التي استخدمت المفهومين السابقين؛ الطريقة الجينية للتصنيف (Genetic Approach to Classification)، حيث أنها استخدمت مفهوم البرنامج الجيني (Genetic Program) لتقديم مفهوم التصنيف المعنوي (Semantic Classification).

في هذا البحث سنقوم بتعديل وتحسين المفاهيم التي استخدمت في الطريقة السابقة لبناء هرمية مجموعة عامة (Generic Class Hierarchy) والتعديل على مفهوم البرنامج الجيني، وذلك لإنشاء مجموعة من هرميات المجموعة الصلبة (Concrete Class Hierarchies) بعد تطبيق مفهوم الوراثة الانتقائية (Selective Inheritance) عليها، ومن ثم سيتم التعامل مع هذه الهرميات لإنشاء مجموعة من الكائنات (Objects) ليتم تخزينها كما هي في قواعد بيانات شبيهة (Object-Oriented Database) بحيث تكون خالية من القيم عديمة الوجود.