

ОНТОЛОГИЧЕСКИ-ОРИЕНТИРОВАННЫЙ ПОДХОД К ПРОБЛЕМЕ ОТКРЫТОСТИ ПРОГРАММНЫХ СИСТЕМ

С.С. Курбатов, А.П. Лобзин, Г.К. Хахалин

Научно Исследовательский Центр Электронной Вычислительной Техники,
Москва, Россия

Введение

В широком смысле *компьютерной открытой системой* целесообразно называть любую систему, архитектура которой соответствует открытым спецификациям. Под открытыми спецификациями подразумеваются опубликованные и общедоступные формализованные описания программно-аппаратных компонентов, предполагающие их взаимодействие с другими компонентами. Именно открытые спецификации позволяют создавать программно-аппаратные комплексы из продуктов разных производителей.

Известно, что полная открытость является недостижимым идеалом для реальных систем, в которых определению открытости соответствуют лишь некоторые части. Понятие открытости достаточно многогранно и не стандартизовано. Открытость можно рассматривать на разных уровнях иерархии программного и аппаратного обеспечения системы или ее составных частей [1]. Средства достижения открытости обобщенно включают: протоколы промышленных сетей, аппаратные интерфейсы (типа RS-232), программные интерфейсы (типа DCOM-технологии), интерфейсы пользователя и программная совместимость.

В данной статье предлагается частное решение проблемы открытости реальной программной системы на основе онтологического подхода. Программная система включает реляционную СУБД (более 250 таблиц) и программный интерфейс (более 2000 программных модулей). Отметим, что

средства аппаратно-программного уровня (протоколы, интероперабельность, сервис-ориентированная архитектура – СОА и т.д.) получили большое развитие. В то же время средства более высокого концептуального уровня (типа баз знаний и онтологий) разработаны гораздо слабее. Однако перспективность онтологического подхода к решению проблемы открытости программной системы, по мнению авторов, является весьма значительной.

Концепция решения

Описываемая ниже логика решения является составной частью целостной системы, объединяющей лингвистическую модель предметной области, ее онтологическое представление и систему генерации из онтологии программного кода по естественно-языковому описанию. Целостная система разрабатывается на уровне фундаментального исследования, основная концепция которого еще далека от завершения, в особенности в части лингвистического компонента. Однако вопросы описания в онтологии схемы и содержания реляционной СУБД и программных модулей интерфейса с базой были логическом уровне разработаны с детальностью, позволяющей воплотить их в макете программной системы.

Разработанная логика представления в онтологии схемы и содержания базы данных включала как алгоритмически простую обработку текстового представления схемы реляционной базы, так и представление более сложной информации об объектах предметной области, их свойствах и отношениях. Концептами в онтологии выступали описания реальных объектов предметной области, их отображений в таблицах реляционной базы, отношений между концептами и их представление в базе (на уровне первичных и вторичных ключей). Таким образом, онтологическое представление является более богатым, чем стандартная ERWIN-схема.

Базовая логика представления программных модулей основной (или объектной) системы включала представление каждого модуля как концепта и анализ текстов программ для выявления свойств и отношений модулей

(вызовы, описания входных и выходных параметров, соответствие стандартам программирования и т.д.). Важным свойством модуля является его естественно-языковое описание, исходной информацией для которого может быть не только ручной ввод в онтологию, но и комментариев в тексте модуля, наименование пункта меню для вызова модуля и т.д.

Предполагается, что в процессе дальнейшего развития лингвистического компонента онтология будет пополняться путем интеллектуального сканирования естественно-языкового текста технического задания на реальную систему. Стиль такого пополнения для текстов энциклопедического характера намечен в [2].

Реализация

Для экспериментальной проверки разработанной логики была выбрана реальная коммерческая система управления сетью торговых предприятий. Основные концепты объектов предметной области были подготовлены средствами редактирования онтологии с элементами лингвистической трансляции текста общего технического задания на всю систему. Далее были разработаны модули анализа схемы и содержания базы данных, заполняющие основные концепты онтологии по текстовым представлениям и обращениям к базе.

Свойства концептов в онтологии формируются как из схемы базы данных, так и из семантики реальных объектов внешнего мира. Например, концепт "касса" может не содержать в схеме базы атрибута "производитель", но в расширенное онтологическое представление целесообразно включить такую информацию. Если в дальнейшем потребуется модернизировать схему базы, онтологическое представление обеспечит удобные возможности для автоматизации такой модернизации.

Отношения между концептами включают как отношения, фиксируемые ERWIN-схемами (первичный/вторичный ключ), так и более семантически нагруженные отношения. Классический пример такого отношения "фирма"

– ”поставляет” - ”товар”, реализуемое в базе через соответствующие ключи в таблицах ”фирма”, ”заголовок поставки”, ”спецификация поставки” и ”каталог товаров”. Отношения типа ”в модуле используются поле F_i из таблицы T_j ” заносятся в онтологию только после анализа текстов программ.

Аналоги онтологических структур, описывающие базу данных на высоком понятийном уровне, разрабатывались достаточно давно – это так называемые концептуальные схемы баз данных. Проблема представления в онтологии программ более сложна и менее разработана. На данном этапе реализации в онтологию были автоматически занесены:

- граф вызовов программ (анализ на оператор RUN);
- таблицы и поля базы данных, используемые в данной программе (анализ текста модуля на операторы обращения к базе данных);
- естественно-языковые описания функционала системы (анализ пунктов меню в головном модуле системы);
- естественно-языковые описания базового функционала системы (анализ основных библиотек системы).

Механизм формирования программ из онтологических структур реализован лишь в первом приближении и является обобщением алгоритмов генерации программного кода из онтологии, которые ранее обрабатывались в задачах формирования SQL-запросов и графических объектов из описания на предметно-ориентированном естественном языке [3].

Обсуждение результатов

Результаты экспериментов с реализованным макетом включали анализ текстов базовой программной системы с использованием онтологии и исследование возможностей автоматической генерации программных кодов из онтологического описания. Анализ позволил автоматически выявить ряд несоответствий текстов (программных кодов) стандартам системы, зафиксированных в правилах онтологии (для данной предметной области).

Это дало возможность быстро улучшить тексты программистами, обладающими весьма слабой квалификацией.

Автоматическая генерация программных кодов из онтологии потребовала редактирования сформированной программы, но в небольшом объеме (менее 5%). Генерация программных кодов на SQL (а не на 4GL) позволило упростить миграцию программ для функционирования на другой базе (в экспериментах использовалась ORACLE). Использование онтологии, обеспечивающей высокую наглядность представления графа вызовов программ и естественно-языковое описание их семантики, существенно упрощает модернизацию системы.

Заключение

В целом экспериментальные результаты подтверждают перспективность использования онтологического подхода с использованием средств лингвистической трансляции для повышения степени открытости реальной системы, включающую реляционную базу данных и программы интерфейса пользователя.

Литература

1. Общие принципы построения вычислительных сетей, [Электронный ресурс]. – Режим доступа: <http://compnets.narod.ru/1-7.html> (дата обращения 29.04.13).
2. Курбатов С.С., Инструментальные средства для автоматизированного формирования баз знаний, Сб. трудов V Международной научно-практической конференции "Интегрированные модели и мягкие вычисления в искусственном интеллекте", т.2, 2009.
3. Литвинович А.В., Система синтеза изображений по тексту на естественном языке // Динамика сложных систем — XXI век, № 1, 2013, С. 65-68.