

ФОРМАЛЬНАЯ СЕМАНТИКА КАНОНИЧЕСКОЙ ИНФОРМАЦИОННОЙ МОДЕЛИ В КОМПОЗИЦИОННОЙ ИНФРАСТРУКТУРЕ РАСПРЕДЕЛЕННЫХ ИНФОРМАЦИОННЫХ СИСТЕМ

С.А. Ступников

Институт проблем информатики РАН [ssa@ipi.ac.ru]

Л.А. Калиниченко

Институт проблем информатики РАН [leonidk@ipi.ac.ru]

1. Введение

В настоящее время наблюдается устойчивая тенденция к все большему вовлечению формальных методов в процесс разработки информационных систем (ИС). Это связано с увеличением количества задач, для решения которых необходимо доказательное рассуждение о свойствах систем. Именно для решения такого рода задач и предназначены опирающиеся на математику формальные методы, воплощаемые в совокупности языков, технологий и инструментальных средств спецификации и верификации сложных ИС.

Основу методов верификации сложных ИС составляет *доказательство теорем* (theorem proving). Система при этом описывается на языке некоторой формальной логики, а требуемые свойства системы задаются как формулы логики. Доказательство теорем представляет собой процесс вывода заданных формул из аксиом логики с использованием правил вывода. Выводы могут быть получены автоматически, либо при помощи средств интерактивного доказательства. В контексте метода доказательства теорем формализуется одно из наиболее важных понятий в области формальных методов конструирования ИС - понятие *уточнения*. Говорят, что система *B уточняет* систему *A*, если пользователь может использовать систему *B* вместо системы *A*, не замечая факта замены *A* на *B*.

При разработке ИС приходится решать разнообразные задачи, интенсивно использующие отношение уточнения, которое должно быть формально доказано. В множестве таких задач можно выделить три больших класса: (1) задачи, связанные с удовлетворением специфических нефункциональных требований к ИС (например, создание *safety-critical или fault-tolerant* систем), (2) задачи интеграции множественных неоднородных источников данных и сервисов, (3) задачи композиции ИС из существующих программных и информационных компонентов в интероперабельных средах, таких как Web services, Grid и различные виды промежуточного слоя, расположенного между операционными системами и прикладными системами. Эти задачи требуют решения ряда семантических проблем, включая:

- синтез канонических моделей для посредников над разнородными источниками информации [3];
- согласование онтологий посредника и информационных источников;
- регистрация неоднородных источников информации в предметных посредниках, требующая преодоления различия информационных контекстов.

Настоящая работа представляет собой формальную базу для решения вышеперечисленных задач на основе метода доказательства теорем.

2. Каноническая информационная модель и необходимость определения ее формальной семантики

Для решения задач интеграции неоднородных информационных источников и задач композиции ИС, необходима каноническая информационная модель. Такая модель нужна для однородного представления разнообразных моделей представления информации, используемых в неоднородных источниках. В качестве канонической модели в данной работе выбран язык СИНТЕЗ [2]. Этот язык содержит средства представления слабоструктурированной информации (фреймы); унифицированную систему типов (конструктор абстрактных типов данных, широкий набор встроенных типов); классы, как множества объектов заданных типов; средства представления потоков работ; язык логических формул для выражения инвариантов типов, запросов к информационным ресурсам и спецификации функций. Определение формальной семантики канонической модели необходимо для проведения доказательных рассуждений о моделях информационных ресурсов, например, рассуждений об уточнении или отображении моделей.

Решение перечисленных задач существенным образом опирается на доказательство отношения уточнения спецификаций систем. Для этого необходима технология и инструментальные средства, поддерживающие формальное доказательство свойства уточнения. В данной работе в качестве таких средств используются *B-технология* (*B-technology*) [1] и *B-Toolkit*. Рабочим языком B-технологии является Нотация Абстрактных Машин (*Abstract Machine Notation, AMN*). Для того, чтобы использовать B-технология для доказательства уточнения спецификаций канонической модели, необходимо корректное отображение канонической модели в AMN.

3. Основные результаты работы

В рамках работы определена комбинированная денотационно-аксиоматическая семантика канонической информационной модели (языка СИНТЕЗ). Определен абстрактный синтаксис канонической модели, семантические домены и семантические функции, отображающие абстрактный синтаксис в семантические объекты.

Построено корректное отображение канонической модели в AMN. Разработаны алгоритмы отображения, доказана их корректность в соответствии с семантикой канонической модели.

Реализован прототип программного средства, осуществляющего автоматическое отображение спецификаций канонической модели в спецификации AMN. Разработанные методы и средства в совокупности представляют собой формальную базу для автоматизированного решения задач разработки ИС трех перечисленных классов.

Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 03-01-00821).

Литература

1. *J.-R. Abrial*. The B-Book. - Cambridge University Press, 1996.
2. *Калиниченко Л.А.* СИНТЕЗ: язык определения, проектирования и программирования интероперабельных сред неоднородных информационных ресурсов. - Москва: ИПИ РАН, 1993.
3. *Л.А. Калиниченко, С.А. Ступников, Н.А. Земцов*. Синтез канонических моделей для интеграции неоднородных источников информации. Москва, ИПИ РАН, 2005.