

ПРЕДСТАВЛЕНИЕ ОНТОЛОГИЧЕСКИХ СПЕЦИФИКАЦИЙ ПОСРЕДНИКА В UML

Брюхов Д.О., Тюрин И.Н.

Институт Проблем Информатики Российской Академии Наук

117333, Российская Федерация, Москва, ул. Вавилова 44-2

e-mail: {brd, turin}@ipi.ac.ru

Данная статья¹ посвящена представлению в UML спецификаций на онтологическом языке посредника (MOL), используемом для описания онтологических понятий. В данной статье рассматривается представление в UML спецификаций, заданных на расширении MOL, эквивалентном языку описания онтологий DAML+OIL.

REPRESENTATION OF MEDIATOR'S ONTOLOGICAL SPECIFICATIONS IN UML

Briukhov D.O., Tyurin I.N.

Institute for Problems of Informatics RAS, Moscow, Russia

e-mail: {brd,turin}@ipi.ac.ru

This paper shows the representations the specifications of Mediator Ontology Language (MOL) in UML notation. The MOL is an extensible language. It contains the core and extensions for each ontology language. In the paper we consider the MOL extension for DAML+OIL. DAML+OIL is proposed by W3C as a standart of ontology language for Web. It is an extension of RDF and RDF Schema languages. MOL is based on the SYNTHESIS language. To visualise the SYNTHESIS specifications (including ontology specifications) we use UML. UML specification is also used for loading the ontology specifications into the mediator's metainformation repository.

Введение

В лаборатории композиционных методов проектирования информационных систем ИПИ РАН разрабатывается архитектура посредника неоднородных информационных коллекций [4], который позволяет работать с распределенными неоднородными коллекциями данных как с интегрированной коллекцией информации.

Посредник поддерживает процесс систематической регистрации и классификации коллекций, содержит унифицированные онтологические

¹ Данная работа выполняется в рамках гранта РФФИ №01-07-90084.

данные и метаинформацию для улучшения обнаружения и композиции существующих коллекций [1].

Регистрация коллекции – это процесс взаимодействия провайдера цифровой коллекции с предметным посредником во время операционной фазы посредника, заключающийся в контекстуализации коллекции в посреднике (согласовании ее понятийного и терминологического контекста с контекстом посредника), представления классов коллекции как материализованных взглядов над классами посредника, генерации адаптеров.

Онтологический язык посредника (MOL [5]) основан на модели языка Синтез [3]. Язык Синтез используется в качестве канонической модели данных посредника неоднородных информационных коллекций. В рамках посредника определяется метаинформация предметной области, которую представляет данный посредник. Одной из составляющих метаинформации является набор онтологических спецификаций предметной области.

В качестве инструмента для наглядной (визуальной) работы с Синтез спецификациями (в частности, онтологическими) используется язык UML [6]. UML спецификации также используются для загрузки спецификаций в базу метаинформации посредника при регистрации новых коллекций.

MOL является расширяемым языком. Он состоит из ядра и расширений, которые строятся для конкретных онтологических языков. Для загрузки онтологических спецификаций в базу метаинформации требуется определить представление онтологических спецификаций на расширении MOL в UML.

В данной статье рассматривается представление в UML онтологических спецификаций, описанных на расширении MOL, эквивалентном DAML+OIL [2]. DAML+OIL предложен W3C в качестве стандарта онтологического языка для Web. Этот язык является расширением языков RDF и RDF Schema и использует систему встроенных типов XML Schema.

Расширение MOL, соответствующее DAML+OIL

Отображение DAML+OIL в MOL задается при помощи следующего набора соответствий:

- классы DAML+OIL отображаются в типы MOL;
- отношение подкласс отображается в отношение тип-подтип;
- свойства типов данных отображаются в атрибуты типов;
- объектные свойства отображаются в атрибуты типов. Данные атрибуты определяются как экземпляры соответствующих метатипов ассоциаций;
- ограничения свойств отображаются в инварианты соответствующих метатипов ассоциаций;

- ограничения кардинальности свойств отображается атрибутом `association_type` соответствующего метатипа ассоциаций;
- встроенные типы данных DAML+OIL (XML Schema) отображаются во встроенные типы данных MOL (Синтез)

Полное описание отображения DAML+OIL в MOL определено в [5]. Ниже приведен пример отображения спецификации DAML+OIL в MOL:

Спецификация на языке DAML+OIL:

```
<daml:Class rdf:ID="Animal">
  <rdfs:subClassOf>
    <daml:Restriction daml:cardinality="2">
      <daml:onProperty rdf:resource="#hasParent"/>
    </daml:Restriction>
  </rdfs:subClassOf>
  <rdfs:subClassOf>
    <daml:Restriction daml:cardinality="1">
      <daml:onProperty rdf:resource="#age"/>
    </daml:Restriction>
  </rdfs:subClassOf>
</daml:Class>

<daml:Class rdf:ID="Male">
  <rdfs:subClassOf rdf:resource="#Animal" />
</daml:Class>

<daml:ObjectProperty rdf:ID="hasParent">
  <rdfs:domain rdf:resource="#Animal" />
  <rdfs:range rdf:resource="#Animal" />
</daml:ObjectProperty>

<daml:ObjectProperty rdf:ID="hasFather">
  <rdfs:subPropertyOf rdf:resource="#hasParent" />
  <rdfs:range rdf:resource="#Male" />
</daml:ObjectProperty>

<daml:DatatypeProperty rdf:ID="age">
  <rdfs:range
rdf:resource="http://www.w3.org/2000/10/XMLSchema#Integer
"/>
</daml:DatatypeProperty>
```

Та же спецификация на языке MOL:

```
{Animal;
  in: type, daml_oil;
  hasParent: Animal;
  metaslot
```

```

        in: HasParent
    end
    hasFather: Male;
    metaslot
        in: HasFather
    end
    age: integer;
}

{HasParent;
  in: metatype, association, daml_oil;
  inverse: HasChild;
  instance_section: {
    domain: Animal;
    range: Animal;
    association_type: {{2,2},{0,inf}}
  }
}

{Male;
  in: type, daml_oil;
  supertype: Animal
}

{HasFather;
  in: metatype, association, daml_oil;
  superclass: HasParent;
  instance_section: {
    domain: Animal;
    range: Male;
  }
}

```

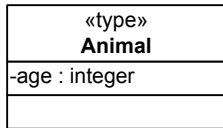
Представление понятий, заданных на языке MOL, в UML

Для отображения онтологий мы используем диаграммы классов (class diagrams) UML.

Для конструкций MOL, для которых существуют эквивалентные конструкции в UML, мы используем соответствующую нотацию UML. В противном случае мы используем механизм стереотипов, который введен в UML для расширения языка. Стереотип представляет собой новый класс моделируемых элементов, который вводится уже во время моделирования. Посредством стереотипов в UML представлена метаклассификация элементов.

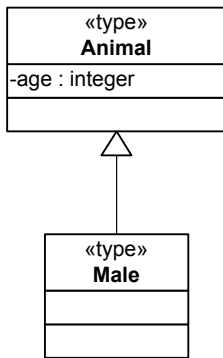
Отображение типов

Типы MOL изображаются классами UML со стереотипом «type»:



Отображение отношения тип/подтип

Отношение тип/подтип изображается посредством отношения обобщения UML:



Отображение атрибутов и ассоциаций

Атрибуты, типами которых являются встроенные типы данных, отображаются в UML как атрибуты классов, а атрибуты, типами которых являются абстрактные типы данных, отображаются в UML как ассоциации.

Атрибуты типа изображаются текстовой строкой вида:

<имя атрибута> : <тип атрибута> = <начальное значение>

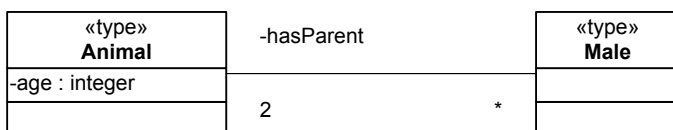
где *<тип атрибута>* - спецификация типа атрибута на языке MOI,

<начальное значение> - выражение, задающее начальное значение атрибута на языке MOI.

Ассоциации представляют семантическую связь между экземплярами типов. Каждый конец ассоциации, где она соединяется с типом называется ролью. Роли содержат основную информацию ассоциации:

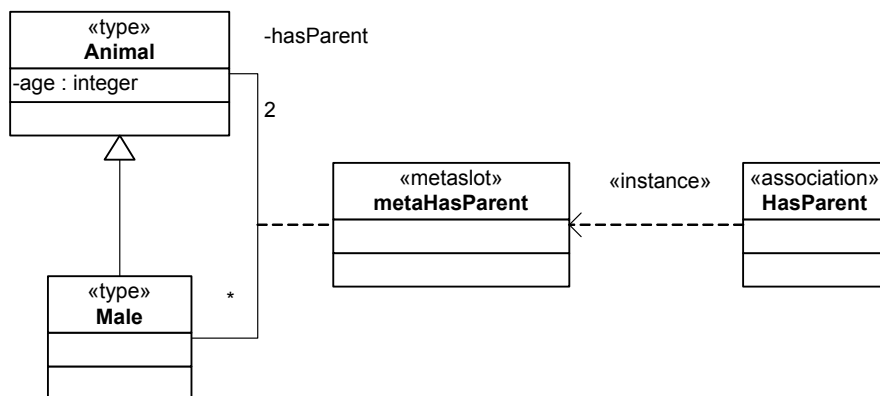
- *имя роли* показывает, как данный тип виден из другого типа;
- *кардинальность* определяет количество экземпляров типа принимающих участие в ассоциации;
- *индикатор агрегации* показывает, что один тип является частью второго типа.

Пример представления атрибута *age* и ассоциации *hasParent* в UML:



Отображение метатипов

Метатипы ассоциаций MOL изображаются классами UML со стереотипом “association”, метаслоты изображаются классами UML со стереотипом “metaslot”:



Заключение

В данной статье мы рассмотрели представление в UML онтологических спецификаций, описанных на расширении ядра MOL, эквивалентном DAML+OIL, который предложен W3C в качестве стандарта онтологического языка для Web.

Литература

- [1] Briukhov, D.O., Kalinichenko, L.A., Skvortsov, N.A. Information sources registration at a subject mediator as compositional development. In Proceedings of the Fifth East European Symposium on Advances in Databases and Information Systems (ADBIS'01), Springer-Verlag, 2001, pp. 70-83.
- [2] DAML+OIL (March 2001) reference description.
<http://www.w3.org/TR/daml+oil-reference>
- [3] Kalinichenko L.A. SYNTHESIS: the language for description, design and programming of the heterogeneous interoperable information resource environment, Moscow, 1995.
- [4] Kalinichenko L.A., Briukhov D.O., Skvortsov N.A., Zakharov V.N. Infrastructure of the subject mediating environment aiming at semantic interoperability of heterogeneous digital library collections. 2nd Russian Conference "DIGITAL LIBRARIES: Advanced Methods and Technologies, Digital Collections", September 26-28, 2000, Protvino
- [5] Kalinichenko L.A., Skvortsov N.A. Extensible Ontological Modeling Framework for Subject Mediation, submitted to ECDL'2002
- [6] Unified Modeling Language (UML) Specification, Version 1.4.
<http://www.omg.org/technology/documents/formal/uml.htm>