

ИНТЕГРАЦИЯ СЕМЕЙСТВА РАСШИРЕННЫХ ДЕСКРИПТИВНЫХ ЛОГИК С РЕЛЯЦИОННОЙ МОДЕЛЬЮ ДАННЫХ

В.А. Резниченко, И.С. Чистякова

Работа является логическим продолжением ранее опубликованного исследования, посвященного созданию способов манипулирования данными. На основе созданной ранее бинарной реляционной структуры данных, осуществляется отображение расширений дескриптивной логики ALC в реляционную модель данных (RDM). В работе используются полученные ранее результаты исследований, а именно структура данных RM^2 , отображения базовых концептов логики ALC в RDM.

Ключевые слова: семантический веб, отображения дескриптивной логики, ALC, RDM, дескриптивная логика, реляционная модель данных, расширения дескриптивной логики, бинарная реляционная структура данных, RM^2 .

Робота є логічним продовженням раніше опублікованого дослідження, яке було присвячено створенню засобів маніпулювання даними. За допомогою раніше створеної бінарної реляційної структури даних здійснюється відображення розширень дескриптивної логіки ALC у реляційну модель даних (RDM). В роботі використовуються результати, що були попередньо отримані, а саме – структура даних RM^2 , відображення базових концептів логіки ALC у RDM.

Ключові слова: семантичний веб, відображення дескриптивної логіки, ALC, RDM, дескриптивна логіка, реляційна модель даних, розширення дескриптивної логіки, бінарна реляційна структура даних, RM^2 .

The paper is a logical continuation of the previously published work, which was dedicated to the creation of the data manipulation methods. Based on the previously created binary relational data structure we perform mappings of the ALC extension into relational data model (RDM). The results of previous research namely data structure RM^2 and mappings of the basic ALC concepts into RDM was used in this paper.

Key words: semantic web, description logic mapping, ALC, RDM, relational data model, description logic extensions, binary relational structure, RM^2 .

Введение

Данная работа является логическим продолжением исследований в области создания механизмов отображения между дескриптивными логиками и реляционной моделью данных. В работе [1] был приведен детальный обзор комплексной проблемы интеграции данных в семантическом вебе. Она состоит из трех основных компонент: выработка схем интеграции данных, выработка отображений между моделями, выработка способов манипулирования данными. Основываясь на результатах этой работы, авторами была проработана вторая составляющая комплексной проблемы интеграции данных, в результате чего был создан механизм отображения базовых концептов дескриптивной логики ALC в реляционную модель данных (RDM), который детальным образом изложен в работе [2].

Данный механизм относится к таким, которые осуществляют взаимодействие между знание-ориентированными системами, и позволяет преодолеть следующие трудности [2]:

- устранение признаков, порождающих гетерогенность данных [1];
- взаимодействие систем, в основе которых лежат классические модели данных, с современными онтолого-ориентированными системами;
- устранение дублирования данных;
- возможность повторного использования различных компонентов системы;
- возможность построения комплексных информационных систем из разнородных составляющих.

Ввиду ряда особенностей комплексной проблемы интеграции данных, приведенных в обзоре [1], в основу механизма отображений легли следующие компоненты:

- центральная схема интеграции данных;
- GAV/LAV представление;
- формально-логический уровень абстракции.

Суть *центральной интеграции данных* подробно изложена в [1]. О сущности *GAV/LAV представлений* рекомендуем ознакомиться в [3]. Под *формально-логическим уровнем абстракции* авторы понимают решение проблемы на уровне моделей, а не на уровне прикладных реализаций. Мы рассматриваем создание механизмов отображений не на уровне прикладных реализаций (онтологий, баз данных и пр.), а на уровне моделей, лежащих в их основе.

Механизм отображения [2], который применяем в текущем исследовании, базируется на трех этих особенностях. В качестве центральной схемы интеграции он использует бинарную реляционную структуру данных, в основе которой лежит дескриптивная логика ALC. Согласно GAV-представлению, механизм должен иметь схему источника. В текущей работе мы продолжаем использовать в этой роли реляционную

© В.А. Резниченко, И.С. Чистякова, 2016

модель данных. Также не выходим за рамки GAV-представлений, оставляя создание отображений на основе LAV-представлений предметом дальнейших исследований.

В разделе 1 кратко изложена суть упоминаемого механизма, с помощью которого проводятся исследования, результаты которых мы представляем в текущей работе. В разделе 2 рассматривается создание отображений расширенной логики ALC в RDM. Раздел 3 посвящен основным выводам.

Интеграция базовых концептов DL ALC в RDM

Поскольку дескриптивная логика (DL) призвана решать совершенно иные задачи, нежели модель данных, постольку был сформулирован тезис: DL можно рассматривать в качестве модели данных. Обоснование этого утверждения приведено в [2]. Проанализировав ряд публикаций, посвященных созданию отображений между реляционной моделью данных и онтологией, мы пришли к выводу, что описанный нами подход имеет ряд преимуществ над прочими. Прежде всего, следует отметить, что независимо от вариантов установления отображения следует отслеживать и управлять следующими ситуациями:

- возможная потеря данных. Преобразование должно быть таким, что результирующая модель содержала в себе все данные исходной модели;
- возможная потеря семантики. Возможна ситуация, когда данные не теряются, но их семантика при преобразовании «теряется». К таким семантическим характеристикам, например, в реляционной модели относятся ограничения целостности. При преобразовании следует либо гарантировать сохранение семантики, либо указывать, в каких случаях какая именно семантика теряется;
- полнота преобразования. Полнота предполагает, что все, что представимо в исходной модели, может быть преобразовано в результирующую модель;
- взаимная однозначность. Она предполагает, что результат преобразования при обратном отображении дает исходный объект исходной модели.

Существует следующая классификация направлений исследований и разработок в области описания отображений между онтологической и реляционной моделью данных [4, 5].

1. Использование промежуточной модели данных. Именно этот подход подробно изложен в фундаментальной работе [2]. Его идея заключается в том, что преобразование производится не непосредственно, а с использованием промежуточной модели, предполагая, что описание отображений между такой моделью-посредником, с одной стороны, и онтологией и реляционной моделью, с другой стороны, производится намного легче и естественнее. В зависимости от исследования, в качестве центральной схемы предлагается использовать UML, ER и др.

2. Непосредственное отображение онтологии в реляционную базу данных. Суть этого подхода заключается в том, что предлагается алгоритм преобразования онтологии в базу данных. Исследователи рассматривают различные модели баз данных для такого преобразования – реляционную, объектную, объектно-реляционную. Смысл всегда один и тот же: набор правил, согласно которым каждый конкретный элемент онтологии преобразуется в элемент реляционной базы данных (RDB). Авторы таких методологий отмечают серьезные трудности, возникающие при попытке сохранить семантику данных онтологии.

3. Непосредственное отображение реляционной базы данных в онтологию. Это наиболее популярное направление среди всех, имеющих отношение к установлению взаимосвязи между онтологиями и RDB. Его суть не сильно отличается от предыдущего пункта. Каждый метод представляет собой некоторые правила, согласно которым каждый элемент базы данных преобразовывается в элемент онтологии. В зависимости от исследования выделяются три направления: отображения схемы (метаданных) RDB в онтологию, отображения с учетом анализа данных БД, отображения с учетом анализа запросов к БД.

4. Создание и использование онтолого-ориентированной базы данных. Суть заключается в создании специализированных БД, которые в качестве данных будут содержать онтологию и её данные.

Приведем здесь краткие выводы, которые мы сформулировали в результате анализа текущих исследований [2], раскрывая суть и основные преимущества нашего метода.

1. Отсутствие формального подхода. В качестве информационной модели семантического веба выбирается онтология и, как правило, в ее лице берется OWL. В связи с тем, что язык OWL сам по себе не является формальным языком, описание механизма отображений дается на содержательном уровне в виде набора описательных правил. В нашем случае для описания отображений мы берем DL, что дает возможность поставить проблему отображений на более формальный уровень.

2. Односторонний механизм описания отображений. Как правило, независимо описываются отображения из онтологии в RDM и наоборот, проблема установления взаимно-однозначного соответствия при этом не решается. Мы стремимся к тому, чтобы в нашем подходе эта проблема была решена.

3. Отображаются только структурные составляющие моделей. Существующие механизмы описывают отображение только структурных компонент модели данных. Абсолютно ничего не говорится об отображении конструкторов концептов и ролей, с одной стороны, и операций реляционной алгебры – с другой. Наш подход

заключается в том, чтобы также отразить и эту составную часть отображений.

4. Конструкции, которые не отображаются. Совершенно не понятно, каким образом отображать в классическую реляционную модель аксиомы DL (равенства, включения и непересекаемости концептов). Мы предлагаем реляционную структуру, которая предоставляет такие возможности.

5. Гипотезы открытого мира и не уникальности имен. DL строится в предположении существования гипотез открытого мира и не уникальности имен. В свою очередь, реляционная модель формулируется из предположения существования гипотез замкнутого мира и уникальности имен. Предлагая свой вариант реляционной структуры, мы отчасти решаем эти противоречия.

Как было сказано, основополагающей частью механизма отображений является центральная схема, в роли которой выступает DL ALC, представленная в виде модели данных. В работе [2] приводится обоснование такого выбора. Для текущего исследования приведем лишь наиболее важные аспекты.

1. **ALC является базой для многих современных DL.** Логика ALCF, ALCN, ALCQ, ALCI, ALC(o), представленные в текущем исследовании, расширяют синтаксис ALC. Это означает, что все базовые понятия ALC обязательно представлены в каждой из этих логик.

2. **ALC представлена в виде модели данных.** Для этого мы создали реляционную структуру, которая является специальным вариантом классической реляционной структуры, в терминах которой мы описали эту модель. Это сделано с помощью языка ER при посредничестве ER-модели. Графическое представление ER-схемы приведено в [2].

3. **Структура не содержит сущностей, которые представляют n-арные отношения.** В структуре присутствуют только те сущности, которые представляют унарные и бинарные отношения, т.к. в DL ALC n-арные отношения и их экземпляры отсутствуют. Следует отметить, что DL с n-арными отношениями существуют, однако они представляют собой специальные расширения ALC, поэтому на данном этапе мы решили отказаться от них.

4. **В нашей структуре представлена как модельная, так и метамодельная часть.** Метамодельная часть – сущности Concept и Role, содержащие перечни имен концептов и ролей соответственно. Модельная часть – Concept Individual и Role Individual, которые представляют собой индивиды концептов и индивиды ролей. В явном виде таких элементов, как индивиды ролей нет в DL, однако мы вводим данную сущность для связывания индивидов концептов с именами ролей.

5. **Наличие сущностей-связок.** С помощью сущностей-связей, которые обозначены как « \Rightarrow », « \Leftarrow » и « \emptyset » для сущностей концептов и ролей, представляются аксиомы равенства, включения и непересекаемости концептов и ролей логики DL ALC. Сущность-связка « \Rightarrow » представлена для индивидов концептов и предназначена для представления аксиом равенства индивидов в DL. Она показывает, что у одного индивида может быть множество имен, то есть не выполняется принцип уникальности. Однако по-прежнему одно имя именуется только один индивид. Следует отметить, что для индивидов ролей мы не вводим такую аксиому, потому что на данный момент этот вопрос недостаточно исследован, и у нас нет никаких оснований утверждать, что одна и та же связь между одними и теми же индивидами может иметь несколько имен. Поэтому, в нашей структуре верно лишь утверждение о том, что экземпляр роли существует и представляет собой пару индивидов концептов, связанную определенной ролью.

6. **Роль существует как самостоятельная единица.** В любой DL роль не может существовать отдельно. В нашей структуре это возможно, и роль представлена её именем, а также именем концепта-домена и именем концепта-диапазона.

7. **Представление бинарной реляционной структуры данных в терминах RDM.** Такую структуру мы обозначили как RM^2 , где индекс 2 показывает наличие не более чем бинарных отношений. Схематическое представление бинарной реляционной структуры данных приведено в [2]. Поскольку мы будем пользоваться понятиями этой структуры, приведем их текстовое описание.

Отношение Concept.

Относится к метамодельной части. Содержит в себе кортежи с именами концептов.

Атрибуты:

ConceptPK – primary key отношения.

ConceptName – имя концепта.

Отношение Role.

Относится к метамодельной части. Содержит в себе кортежи с именами ролей.

Атрибуты:

RolePK – primary key отношения.

RoleName – имя роли.

Отношение ConceptIndividual.

Относится к модельной части. Содержит в себе кортежи с именами индивидов.

Атрибуты:

ConIndPK – primary key отношения.

ConIndName – имя индивида.

Отношение RoleIndividual.

Относится к модельной части. Содержит в себе кортежи, которые представляют собой индивиды ролей. В явном виде таких экземпляров в DL нет, поэтому в нашей структуре это отношение определяется через foreign key на индивид R-последователя, foreign key на индивид R-преемника и primary key, определяющий конкретный экземпляр, соединяющий оба конкретных индивида.

Атрибуты:

RoleIndPK – primary key отношения.

DomainConIndFK – foreign key на индивид R-преемника.

RangeConIndFK – foreign key на индивид R-последователя.

Отношения LinkConceptIndividual и LinkRoleIndividual.

Дополнительные отношения RM^2 . Поскольку один концепт может иметь множество индивидов, а один и тот же индивид может принадлежать нескольким концептам, то для разрешения связи многие-ко-многим мы вводим дополнительное отношение LinkConceptIndividual. Аналогично, отношение LinkRoleIndividual разрешает связь многие-ко-многим между ролями и индивидами, которые она связывает. Каждая роль имеет много экземпляров, связывающих два индивида. С другой стороны, каждая конкретная пара индивидов может быть связана несколькими ролями.

Атрибуты LinkConceptIndividual:

LinkConFK – foreign key на отношение Concept.

LinkConIndFK – foreign key на отношение ConceptIndividual.

Атрибуты LinkRoleIndividual:

LinkRoleFK – foreign key на отношение Role.

LinkRoleIndFK – foreign key на отношение RoleIndividual.

Для описания отображений нам также потребуются операции реляционной алгебры. Их полный перечень вместе с синтаксисом и семантикой приведены в [2].

Для описания отображений нам потребуется синтаксис ALC. Приведем его:

$T \mid \perp \mid A \mid \neg C \mid \text{СПD} \mid \text{СЦD} \mid \exists R.C \mid \forall R.C$

Согласно определениям [6, 7]:

– *синтаксис* DL описывает, какие выражения (концепты, роли, аксиомы и т. п.) считаются правильно построенными в данной логике;

– *семантика* указывает, как интерпретировать синтаксические выражения, т. е. придать им формальный смысл.

На основании описания реляционной структуры RM^2 :

– синтаксические концепты ALC будут считаться правильно построенными в RM^2 ;

– классическая семантика полностью сохраняется в ALC;

– **интерпретация синтаксиса ALC в терминах структуры RM^2 и является искомым набором отображений**, описывающим интерпретацию синтаксиса DL ALC в RDM.

Приведем набор обозначений, необходимых для описания отображений. В таблице приведены все базовые составляющие синтаксиса ALC. Индекс DL обозначает классическое семантическое обозначение, отличное от обозначений отображений в структуре RM^2 . Показатель E (экстенционал) обозначает синтаксические конструкции, которые интерпретируются в RM^2 .

Как уже упоминалось, в DL не существует роль как самостоятельная единица, поэтому для нее не приводится семантическое обозначение в строке ALC. Однако, она существует в описываемой нами реляционной структуре, потому для ее обозначения вводится специальный символ.

Таблиця. Набор обзначений для описания отображений

	C	D	R	$\neg C$	$C \cap D$	$C \sqcup D$	$\exists R.C$	$\forall R.C$
ALC	C_{DL}^I	D_{DL}^I	–	$(\neg C)_{DL}^I$	$(C \cap D)_{DL}^I$	$(C \cup D)_{DL}^I$	$(\exists R.C)_{DL}^I$	$(\forall R.C)_{DL}^I$
RM ²	$C_{RM^2}^E$	$D_{RM^2}^E$	$R_{RM^2}^E$	$(\neg C)_{RM^2}^E$	$(C \cap D)_{RM^2}^E$	$(C \cup D)_{RM^2}^E$	$(\exists R.C)_{RM^2}^E$	$(\forall R.C)_{RM^2}^E$

Приведем отображения всех этих конструкций. Мы не будем углубляться в объяснение, поскольку оно присутствует в [2]. Приведем лишь конечные формулы.

Согласно [7], все выражения в семантике ALC интерпретируются индукцией по построению концепта, с помощью интерпретирующей функции. Следуя этому правилу:

- каждому атомарному концепту A соответствует произвольное подмножество $A^I \subseteq \Delta$;
- каждой атомарной роли R соответствует произвольное подмножество $R^I \subseteq \Delta \times \Delta$;
- концепт истина (Thing) представляет собой всю область определения $T^I \subseteq \Delta$;
- концепт ложь (Nothing) представляет собой пустое множество.

В нашей структуре данные выражения будут интерпретироваться следующим образом:

- каждый концепт (атомарный и обычный) будут представлять собой экземпляр отношения Concept;
- каждая роль будет представлять собой экземпляр отношения Role;
- концепт истина (Thing) представляет собой отношение ConceptIndividual;
- концепт ложь (Nothing) останется пустым множеством.

Отображения концептов C и D.

$$C_{RM^2}^E = \pi_{\text{ConIndName}} (\sigma_{\text{ConceptName}='C'} (\text{ConceptIndividual} \bowtie_{\text{ConIndPK}=\text{LinkConIndFK}} (\text{LinkConceptIndividual} \bowtie_{\text{LinkConFK}=\text{ConceptPK}} \text{Concept}))).$$

Описание концепта D происходит аналогично концепту C, но на шаге 2 селекция производится не по «C», а по «D».

Отображение роли R.

$$R_{RM^2}^E = \pi_{\text{Domain, Range}} (\rho_{\text{ConIndName}/\text{Domain}} (\text{ConceptIndividual} \bowtie_{\text{ContIndPK}=\text{DomainConIndFK}} (\rho_{\text{ConIndName}/\text{Range}} (\text{ConceptIndividual} \bowtie_{\text{ContIndPK}=\text{RangeConIndFK}} (\sigma_{\text{RoleName}='R'} (\text{RoleIndividual} \bowtie_{\text{RoleIndPK}=\text{LinkRoleIndFK}} (\text{LinkRoleIndividual} \bowtie_{\text{LinkRoleFK}=\text{RolePK}} \text{Role})))))))).$$

Отображение дополнения концепта $\neg C$.

$$(\neg C)_{RM^2}^E = \pi_{\text{ConIndName}} (\text{ConceptIndividual}) - C_{RM^2}^E.$$

Отображение объединения концептов (C UD).

$$(C \sqcup D)_{RM^2}^E = C_{RM^2}^E \cup D_{RM^2}^E.$$

Отображение пересечения концептов (C ∩ D).

$$(C \cap D)_{RM^2}^E = C_{RM^2}^E \cap D_{RM^2}^E.$$

Отображение конструктора экзистенциального ограничения ($\exists R.C$).

$$(\exists R.C)_{RM^2}^E = \pi_{\text{Domain}}(R_{RM^2}^E \bowtie_{\text{Range} = \text{ConIndName}} C_{RM^2}^E).$$

Отображение конструктора ограничения значения ($\forall R.C$).

$$(\forall R.C)_{RM^2}^E = \pi_{\text{Domain}} R_{RM^2}^E - \pi_{\text{Domain}}(R_{RM^2}^E \cap (\pi_{\text{Domain}} R_{RM^2}^E \times (\pi_{\text{Range}} R_{RM^2}^E - \pi_{\text{ConIndName}} C_{RM^2}^E))).$$

Отображение расширений DL ALC в RDM

Существуют многочисленные расширения логики ALC путем добавления новых конструкторов для построения концептов и ролей, либо добавления новых видов аксиом. В работе [8] выделяют две категории расширений: «классические» и «не классические». Строго формального определения у категорий нет, есть лишь существенные отличия.

К «классическим» относятся такие расширения, чья семантика может быть интерпретирована с помощью теоретико-модельного подхода, в то время как определение семантики «не классических» конструкторов проблематичнее, и требует расширения теоретико-модельного подхода. Ввиду отсутствия строгой классификации, «классические» расширения можно так же разделить на условные две категории:

- расширения путем добавления конструкторов концептов;
- расширения путем добавления операций над ролями.

К первой категории относятся такие, которые добавляют к стандартному набору конструкторов концептов новые конструкторы, с помощью которых получаются выражения, также являющиеся концептами. Ко второй категории относятся такие расширения, которые вводят в дескриптивную логику операции над ролями.

В текущей публикации мы рассматриваем первую категорию расширений, путем добавлений новых концептов, а также обратные роли, которые относятся ко второй категории расширений операциями над ролями.

Имеется неформальное соглашение об именовании получающихся при этом логик: обычно каждому конструктору или виду аксиом сопоставляют букву, которую (за редкими исключениями) добавляют к имени логики. Например, если логика содержит только базовый синтаксис ALC и качественное ограничение, обозначаемое буквой Q, то такая логика будет называться ALCQ. Если к этой логике добавить инверсию ролей, обозначаемую буквой I, то такая логика будет называться ALCIQ. Это универсальный механизм, которым пользуется большинство исследователей дескриптивных логик.

В нашей статье для каждого расширения мы будем указывать: букву, добавляемую к имени логики; название новой связки (или аксиомы); ее синтаксис и семантику; а также формулу отображения.

Описание некоторых расширений для DL ALC.

Приведем ряд определений, необходимых для описания расширений [7].

Семантика логики ALC задается с помощью интерпретаций $I = (\Delta, \cdot^I)$. Здесь интерпретирующая функция \cdot^I сопоставляет каждому атомарному концепту $A \in CN$ подмножество $A^I \subseteq \Delta$, атомарной роли $R \in RN$ – бинарное отношение $R^I \subseteq \Delta \times \Delta$, а имени индивида $a \in IN$ (если таковые были в языке, например, в ABox) – элемент $a^I \in \Delta$. Введем также удобное обозначение: для элемента $e \in \Delta$ множество его R-последователей будем обозначать как $R^I(e) = \{d \in \Delta \mid \langle e, d \rangle \in R^I\}$. Мощность какого-либо множества M будем обозначать $\#M$, а множество имен индивидов IN .

R-последователь – это такой индивид, который является правой частью роли R.

R-преемник – это такой индивид, который является левой частью роли R.

Следующие конструкторы называются *численными ограничениями на роли*:

- **(F) функциональность**: это концепт, у которого количество R-последователей не превышает 1. Если R – роль, то $(\leq 1R)$ – концепт. Очевидно, что в этом языке, используя отрицание, можно выразить концепт $(\geq 2R)$ как $\neg(\leq 1R)$. Другие ограничения недоступны.

- **(N) количественное ограничение**: это концепт, у которого количество R-последователей $\geq n$ или $\leq n$. Если R – роль, а $n \geq 0$ натуральное число, то $(\leq nR)$ и $(\geq nR)$ – концепты. На основании отношений \geq и $=$ с помощью операции пересечения можно построить отношение равенства, например, $(\leq n \cap \geq n) \Rightarrow =n$. А на основании \leq , \geq и $=$ можно выразить строгое неравенство, например $(\leq n \cap \neg =n) \Rightarrow <n$.

- **(Q) качественное ограничение**: это концепт, у которого количество R-последователей $\geq n$ или $\leq n$, и все R-последователи принадлежат концепту C. Если R – роль, C концепт, а $n > 0$ натуральное число, то $(\leq n R.C)$ и $(\geq n R.C)$ – концепты. Проведя, аналогичную предыдущему пункту, цепочку логических рассуждений,

становится ясно, что выразимы также ограничения вида ($<n$ R.C) и ($>n$ R.C).

Далее рассмотрим ещё два конструктора, рассмотренные в текущей работе, первое из которых относится к операциям над ролями, а второе – к «расширениям концептами», но не входит в группу численных ограничений.

- **(I) обратные роли:** если R – атомарная роль, то R – является ролью. Обратные роли можно использовать там, где использовались обычные (атомарные) роли, т. е. в выражениях с кванторами, в численных ограничениях на роли (если они доступны в языке), а также в аксиомах для ролей.

- **(O) номиналы:** если a – имя индивида (т. е. $a \in IN$), то {a} – концепт. В текущей публикации мы не рассматриваем отображение для данного концепта.

Перейдем к семантике приведенных концептов и описанию их отображений.

Отображения логик ALCF, ALCN, ALCQ в RDM.

Исходя из этих определений семантики расширений F, N, Q, приведенных выше, мы можем сказать, что семантика численных ограничений одинакова, и представлена следующим образом:

$$(\leq nR)^I = \{e \in \Delta \mid \#R^I(e) \leq n\},$$

$$(\leq nR.C)^I = \{e \in \Delta \mid \#(R^I(e) \cap C^I) \leq n\}.$$

Для наглядности, рассмотрим каждое ограничение по отдельности, с применением обозначений таблицы.

Семантика **функциональности** имеет следующий вид:

$$(\leq 1R)^I_{DL} = \{e \in \Delta \mid \#R^I(e) \leq 1\},$$

$$(\geq 2R)^I = \{e \in \Delta \mid \#R^I(e) \geq 2\}.$$

Как было сказано, другие ограничения недоступны.

Если мы найдем отображение для концепта $(\geq 2R)^I_{DL}$, то оно будет справедливо также для выражения $\neg(\leq 1R)^I_{DL}$.

Исходя из отображений, полученных в результате предыдущих исследований, мы знаем, что отображение роли $R^E_{RM^2}$ имеет два столбца Domain и Range.

Перейдем к созданию отображения:

- 1) сначала соединим $R^E_{RM^2}$ самим с собой по условию равенства доменов.

$$(R^E_{RM^2} \bowtie_{\text{Domain}=\text{Domain}} R^E_{RM^2});$$

- 2) Далее сделаем выборку по условию неравенства диапазонов.

$$\sigma_{\text{Range} \neq \text{Range}}(R^E_{RM^2} \bowtie_{\text{Domain}=\text{Domain}} R^E_{RM^2});$$

- 3) Выполняем проекцию по домену.

$$\pi_{\text{Domain}}(\sigma_{\text{Range} \neq \text{Range}}(R^E_{RM^2} \bowtie_{\text{Domain}=\text{Domain}} R^E_{RM^2})).$$

Таким образом, в результирующую выборку попадут те R-преемники, у которых два и более R-последователей.

Итак, отображение для концепта $(\geq 2R)^I$ имеет вид:

$$(\geq 2R)^E_{RM^2} = \pi_{\text{Domain}}(\sigma_{\text{Range} \neq \text{Range}}(R^E_{RM^2} \bowtie_{\text{Domain}=\text{Domain}} R^E_{RM^2})).$$

Для концепта $(\leq 1R)^I_{DL}$ справедливо выражение $\Delta - (\geq 2R)^E_{RM^2}$. Исходя из вышеизложенного, отображение для данного концепта будет иметь следующий вид:

$$(\leq 1R)^E_{RM^2} = \Delta - \pi_{\text{Domain}}(\sigma_{\text{Range} \neq \text{Range}}(R^E_{RM^2} \bowtie_{\text{Domain}=\text{Domain}} R^E_{RM^2})).$$

Перейдем к семантике *количественного ограничения*, которое имеет следующий вид:

$$(\leq nR)_{DL}^I = \{e \in \Delta \mid \#R^I(e) \leq n\},$$

$$(\geq nR)_{DL}^I = \{e \in \Delta \mid \#R^I(e) \geq n\}.$$

Отображения данных концептов являются обобщением функциональности. Мы должны выполнить схожие шаги n раз, за одним исключением. Вместо обычного соединения, на первом шаге выполним естественное соединение (обозначается $*$) – соединение по равенству атрибутов, после которого дубликаты атрибутов удаляются. Покажем это детальнее:

1. Выполним естественное соединение $R_{RM^2}^E$ по условию равенства доменов n раз с предварительным переименованием как самого отношения, так и атрибута Range.

$$\left(\underset{1 \leq i \leq n}{*} \rho_{\text{Ri}(\text{Domain}, \text{Rangei})}(R_{RM^2}^E) \right).$$

Здесь мы выполнили операцию переименования $R_{RM^2}^E$, чтобы отличать порядковый номер i отношения в соединении. Обозначение $\underset{1 \leq i \leq n}{*}$ показывает, что необходимо выполнить n естественных соединений отношения.

2. Далее сделаем выборку по условию попарного неравенства всех диапазонов.

$$\underset{1 \leq i \leq n}{\sigma} \text{ Ri.Rangei} \neq \text{Rj.Rangej} \left(\underset{1 \leq i \leq n}{*} \rho_{\text{Ri}(\text{Domain}, \text{Rangei})}(R_{RM^2}^E) \right).$$

3. Выполняем проекцию по домену.

$$\prod_{\text{Domain}} \underset{1 \leq i < j \leq n}{\sigma} \text{ Ri.Rangei} \neq \text{Rj.Rangej} \left(\underset{1 \leq i \leq n}{*} \rho_{\text{Ri}(\text{Domain}, \text{Rangei})}(R_{RM^2}^E) \right).$$

Итак, мы нашли отображение для концепта $(\geq nR)_{DL}^I$, которое выглядит следующим образом:

$$(\geq nR)_{RM^2}^E = \prod_{\text{Domain}} \underset{1 \leq i < j \leq n}{\sigma} \text{ Ri.Rangei} \neq \text{Rj.Rangej} \left(\underset{1 \leq i \leq n}{*} \rho_{\text{Ri}(\text{Domain}, \text{Rangei})}(R_{RM^2}^E) \right).$$

Исходя из этого, отображение для концепта $(\leq nR)_{DL}^I$ будет иметь следующий вид:

$$(\leq nR)_{RM^2}^E = \Delta - (\geq (n+1)R)_{RM^2}^E.$$

Как было сказано, концепты $(< nR)_{DL}^I$ и $(> nR)_{DL}^I$ можно отобразить $(\geq nR)_{RM^2}^E$, $(\leq nR)_{RM^2}^E$ и $(= nR)_{RM^2}^E$.

Для этого покажем, что такое $(= nR)_{RM^2}^E$:

$$(= nR)_{RM^2}^E = (\geq nR)_{RM^2}^E \cap (\leq nR)_{RM^2}^E.$$

Таким образом, отображения концептов $(< nR)_{DL}^I$ и $(> nR)_{DL}^I$ имеют вид:

$$(> nR)_{RM^2}^E = (\geq nR)_{RM^2}^E - (= nR)_{RM^2}^E,$$

$$(< nR)_{RM^2}^E = (\leq nR)_{RM^2}^E - (= nR)_{RM^2}^E.$$

Перейдем к семантике *качественного ограничения*, имеющего следующий вид:

$$(\leq nR.C)^I = \{e \in \Delta \mid \#(R^I(e) \cap C^I) \leq n\},$$

$$(\geq nR.C)^I = \{e \in \Delta \mid \#(R^I(e) \cap C^I) \geq n\}.$$

Отображення для качественного ограничения похожи на отображения количественного ограничения, с одним уточнением: индивиды диапазона должны принадлежать концепту C , что достигается предварительным соединением $R_{RM^2}^E$ с $C_{RM^2}^E$:

Отбросим из $R_{RM^2}^E$ те кортежи, у которых R-последователи не принадлежат $C_{RM^2}^E$:

$$R_{RM^2}^E \bowtie_{\text{Range} = \text{ConIndName}} C_{RM^2}^E.$$

Далее выполняем точно те шаги, что и для количественного ограничения, применяя их к полученному выше отношению. В результате получим:

$$\Pi_{\text{Domain}} \left(\sigma_{\substack{1 \leq i < j \leq n \\ \text{Ri.Rangei} \neq \text{Rj.Rangej}}} \left(\rho_{\text{Ri}(\text{Domain}, \text{Rangei})} \left(R_{RM^2}^E \bowtie_{\text{Range} = \text{ConIndName}} C_{RM^2}^E \right) \right) \right).$$

Таким образом, отображения для качественного ограничения имеют вид:

$$(\geq nR.C)_{RM^2}^E = \Pi_{\text{Domain}} \left(\sigma_{\substack{1 \leq i < j \leq n \\ \text{Ri.Rangei} \neq \text{Rj.Rangej}}} \left(\rho_{\text{Ri}(\text{Domain}, \text{Rangei})} \left(R_{RM^2}^E \bowtie_{\text{Range} = \text{ConIndName}} C_{RM^2}^E \right) \right) \right),$$

$$(\leq nR.C)_{RM^2}^E = \Delta - (\geq (n+1)R.C)_{RM^2}^E,$$

$$(> nR.C)_{RM^2}^E = (\geq nR.C)_{RM^2}^E - (= nR.C)_{RM^2}^E,$$

$$(< nR.C)_{RM^2}^E = (\leq nR.C)_{RM^2}^E - (= nR.C)_{RM^2}^E.$$

Отображение ALCI в RDM.

Семантика **обратных ролей** выглядит следующим образом:

$$(R^-)_{DL}^1 = \{ \langle e, d \rangle \in \Delta \times \Delta \mid \langle d, e \rangle \in R^1 \}.$$

В бинарном отношении четко разграничены R-последователи и R-преемники. В обратных ролях они меняются местами. В реляционной модели такая операция перестановки выражается путем перестановки-имен атрибутов (сами колонки атрибутов не переставляются).

Исходя из этого, отображение для концепта $(R^-)_{DL}^1$ будет иметь следующий вид:

$$(R^-)_{RM^2}^E = (\rho_{(\text{Range}, \text{Domain})} (R_{RM^2}^E)).$$

Выводы

В данном исследовании, на основании ранее полученных результатов, мы установили отображения для численных ограничений дескриптивной логики ALC, а также для обратных ролей. Таким образом, мы специфицировали механизмы отображений семейства дескриптивных логик, которое основано на добавлении к базовой логике ALC новых конструкторов концептов, а также логику ALCI, которая образуется путем добавления операции обратной роли к базовым конструкторам концептов ALC.

1. Чистякова И.С. Онтолого-ориентированная интеграция данных в семантическом вебе // Проблемы програмування. 2014. – № 2–3. – С. 188 – 196.
2. Резниченко В.А., Чистякова И.С. Отображение дескриптивной логики ALC в бинарную реляционную структуру данных // Проблемы програмування. – 2015. – № 4. – С. 13–30.
3. Lenzerini M. Data Integration: A Theoretical Perspective // Proc. of the 21st ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems (PODS 2002). – N. Y.: ACM Press. – 2002. – P. 233–246.
4. Vysniauskas E., Nemuraite L. Transforming ontology representation from OWL to relational database. Information Technology and Control. – 2006, Vol. 35. N 3A. – P. 333–343.
5. Carmen Martinez-Cruz • Ignacio J. Blanco • M. Amparo Vila. Ontologies versus relational databases: are they so different? A comparison. Artif Intell Rev. – 2012. – 38. – P. 271–290.
6. Baader F., Calvanese D., McGuinness D., Nardi D., and Patel-Schneider P. F., editors. The Description Logic Handbook. Cambridge University Press, 2003.
7. Evgeny Zolin. Description logics (lections). [Online] Available from: <http://lpcs.math.msu.su/~zolin/dl/> [Accessed: 2009].

References

1. CHYSTIAKOVA I.S. (2014). Ontology-oriented data integration on the Semantic Web (Онтолого-ориентированная интеграция данных в семантическом вебе). *Problems in programming*. №2-3, pp.188-196.
2. REZNICHENKO V.A. (2014). Ontology-oriented data integration on the Semantic Web (Онтолого-ориентированная интеграция данных в семантическом вебе). *Problems in programming*. №4, pp. .
3. LENZERINI M. (2002). *Data Integration: A Theoretical Perspective*. Proceedings of the 21st ACM SIGMOD-SIGACT-SIGART symposium on Principles of database systems. New York, NY, USA. 2002. New York: ACM.
4. VYSNIAUSKAS E. et al. (2006). Transforming ontology representation from OWL to relational database. *Information Technology and Control*. №35(3A)., pp.333-343..
5. MARTINEZ-CRUZ C. et al. (2012). Ontologies versus relational databases: are they so different? A comparison. *Artificial Intelligence Review*. №38, pp. 271-290.
6. BAADER F. et al. (2003). *The Description Logic Handbook* pp.47-65, p.95.
7. EVGENY ZOLIN (2009) *Description logics (lections)*. [Online] Available from: <http://lpcs.math.msu.su/~zolin/dl/> [Accessed: 2009].

Об авторах:

Резниченко Валерий Анатольевич,

кандидат физико-математических наук, старший научный сотрудник.

Количество научных публикаций в украинских изданиях – 57.

Количество научных публикаций в иностранных изданиях – 3.

<http://orcid.org/0000-0002-4451-8931>,

Чистякова Инна Сергеевна,

младший научный сотрудник.

Количество научных публикаций в украинских изданиях – 8.

Количество научных публикаций в иностранных изданиях – 0.

<http://orcid.org/0000-0001-7946-3611>.

Место работы авторов:

Институт программных систем НАН Украины
03187, Киев, проспект Академика Глушкова, 40.

Тел.: (044) 526 5139; (044) 526 6249.

E-mail: vreznichenko_47@mail.ru,

inna_islyamova@ukr.net.