

АВТОМАТИЗАЦИЯ ПРОЦЕССОВ ИЗМЕРЕНИЯ ФИЗИКО-ХИМИЧЕСКИХ ВЕЛИЧИН НА БАЗЕ ОНТОЛОГИИ

С. П. КОВАЛЕВ

Институт вычислительных технологий СО РАН, Новосибирск, Россия

e-mail: kovalyov@nsc.ru

Н. А. ПРОКОПОВ

Новосибирский государственный университет, Россия

e-mail: prokopov@ccfit.nsu.ru

This paper describes an approach for application of ontology-based Semantic Web technologies for development of the data acquisition and processing systems. Ontology of measuring systems, based on metrological standards, is proposed. This ontology was applied to the atlas “Atmospheric aerosols of Siberia” that aims to collect atmospheric physico-chemical characteristics over the Siberian region. Some applications of the ontology usage for the design of various measuring systems are considered.

Введение

Задачи измерения физико-химических величин, накопления и обработки результатов измерений возникают практически во всех областях науки и техники. Для их решения активно разрабатываются и применяются автоматизированные информационно-измерительные комплексы, состоящие из специализированных приборов, управляемых программными системами. Техническая база для проведения измерений разрабатывается и систематизируется в рамках науки метрологии, тогда как задачи сбора, хранения и обработки данных относятся к сфере информационных технологий.

При проектировании информационно-измерительных систем необходимо учитывать, что результаты измерений в отрыве от контекста зачастую не имеют большого прикладного значения: интерес представляет их корреляция с другими показателями из смежных областей. Например, изучение влияния распределения атмосферных аэрозолей на состав и состояние растительности в Сибирском регионе может открыть новые, ранее не известные причинно-следственные взаимосвязи. На практике же оказывается, что из-за неверных проектных решений, принятых при создании информационно-измерительных систем, выявление подобных корреляций оказывается весьма трудоемким.

В каждой частной предметной области для хранения и обработки данных традиционно создаются свои собственные узкоспециализированные программные системы,

разрабатываются базы данных, ориентированные на учет отдельного конкретного процесса измерения, а для передачи данных используются закрытые или плохо документированные форматы. Изолированность и разобщенность приводят к тому, что такие системы часто несовместимы (или плохо совместимы) друг с другом, и организация их взаимодействия требует разработки сложного промежуточного программного обеспечения. Глобальная интеграция больших совокупностей таких систем сопряжена со значительными трудностями.

Для повышения эффективности решения задач накопления, обработки, поиска и представления результатов измерений привлекаются современные достижения в области информационных технологий. Особый интерес представляют такие технологии, как XML, RDF и OWL, активно развивающиеся в рамках направления Semantic Web. Они позволяют накапливать, хранить, передавать большие объемы структурированных данных и связанных с ними метаданных унифицированным способом, не зависящим от особенностей реализации отдельных систем. Использование технологий Semantic Web при разработке программных систем измерения и сбора данных позволяет существенно упростить проблему совместимости систем из смежных областей и является первым шагом к построению высокоинтеллектуальных компонентов и агентов [1].

Одним из краеугольных камней подхода Semantic Web является концепция онтологии предметной области — формальной модели основных терминов предметной области, отношений между ними и правил логического вывода предметных знаний. В настоящей статье предлагается вариант онтологии измерительных систем, разработанный на основе метрологических стандартов.

Краткий обзор языка описания онтологий OWL приводится в разд. 1. Раздел 2 посвящен детальному рассмотрению онтологии измерительных систем. О практическом опыте использования этой онтологии рассказывается в разд. 3. Раздел 4 содержит примеры вариантов использования онтологии при проектировании информационно-измерительных систем и систем сбора данных. В заключении сформулированы выводы и направления дальнейших исследований.

1. Язык описания онтологий OWL

В рамках инициативы SemanticWeb 10 февраля 2004 г. консорциум W3 Consortium рекомендовал к использованию стандарт Ontology Web Language (OWL) [2]. Это язык, предназначенный для описания знаний в формате, пригодном для машинного восприятия и обработки. Блоки знаний, оформленные в виде конструкций языка OWL и описывающие некоторую предметную область или ее часть, получили название онтологий.

Основными понятиями, которыми оперирует язык OWL, являются классы, экземпляры, атрибуты и отношения. Классам отвечают основные понятия предметной области, образующие словарь ее терминов. С помощью атрибутов можно описать свойства, присущие тем или иным классам. Отношения задают взаимосвязи между классами. Например, одним из наиболее широко употребляемых отношений является `<is-a>`, означающее, что все экземпляры одного класса должны являться также и экземплярами второго (один из видов наследования). Экземпляры позволяют привязывать введенную классификацию к конкретным объектам. Также язык OWL обладает гибкой системой записи ограничений, с помощью которых для классов, их атрибутов и отношений можно записать практически любое утверждение, формализуемое в логике первого порядка.

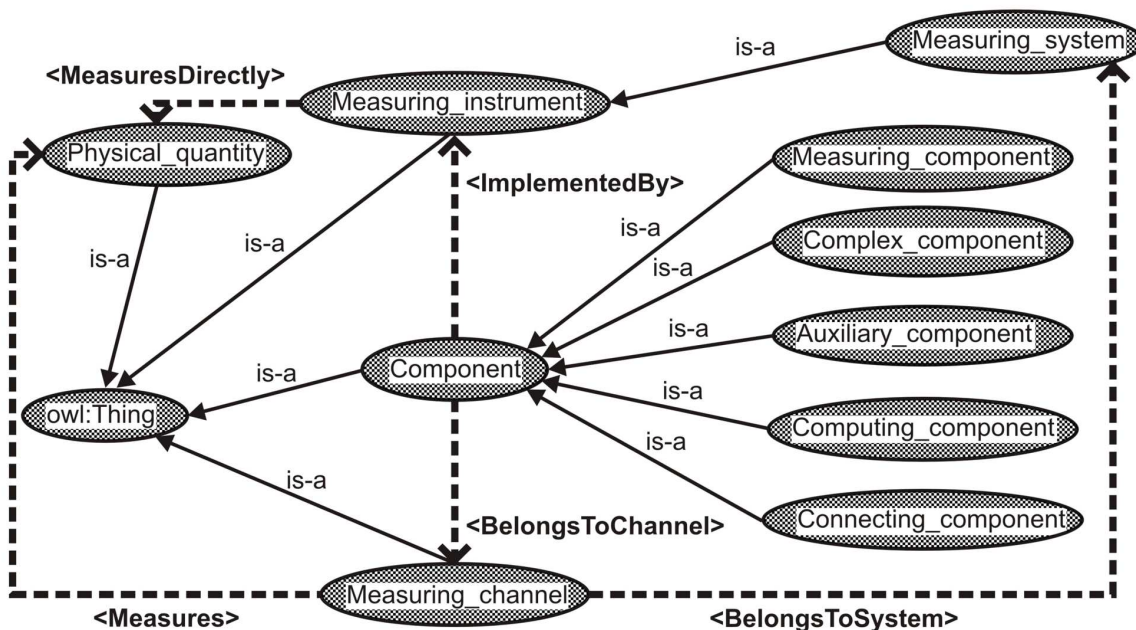
Онтологии могут быть использованы везде, где требуется обработка данных, учитывающая их семантику. В силу изначальной ориентированности языка OWL на машинную обработку правильное применение онтологий может, с одной стороны, существенно упростить и, с другой — открыть новые возможности в разработке приложений, решающих задачи автоматизированной обработки и доступа к данным.

Язык OWL предусматривает возможность совместного использования онтологий: термины из одной онтологии могут содержать ссылки или определяться через термины из другой онтологии. Это позволяет избежать дублирования информации и дает возможность многократно использовать уже разработанные онтологии при создании новых. Так, в онтологии измерительных систем фигурирует понятие *физической величины*, которое может быть задано как ссылка на соответствующий класс из отдельной онтологии физических величин и единиц измерения.

2. Онтология измерительных систем

При измерении любой физической величины важно знать не только сам результат измерения, но и его пространственно-временную привязку, условия, в которых производились измерения, характеристики измерительных приборов и т. п. В том или ином виде любая система, производящая измерения, хранит такие метаданные наряду с результатами измерений.

Процесс измерения физических величин хорошо изучен и регламентирован в рамках науки метрологии. Основные ее положения изложены в стандартах и нормативных документах, поэтому на этапе моделирования предметной области следует опираться именно на них. ГОСТ Р 8.596-2002. Метрологическое обеспечение измерительных систем [3] определяет *измерительную систему* как совокупность компонентов (измерительных, связующих, вычислительных), образующих измерительные каналы. *Измери-*



Онтология измерительных систем.

тельный канал — это логическая сущность, объединяющая весь комплекс измерений и преобразований по получению результата измерения заданной физической величины.

На основе данного метрологического стандарта была разработана онтология измерительных систем. Общая схема этой онтологии представлена на рисунке.

Отметим, что записи элементов онтологии делаются на английском языке. Это позволяет использовать термины из онтологии в коде разрабатываемых программных систем, а также упрощает возможную интеграцию с системами и онтологиями, разрабатываемыми в других странах. Для последующей локализации используются встроенные средства языка XML (который является базовым для OWL), в частности атрибут `xml:lang`. На данный момент онтология содержит названия, определения терминов и комментарии на русском и английском языках.

В онтологии представлены следующие понятия.

Физическая величина (*Physical_quantity*) — объект измерения *измерительной системы*. Это может быть мгновенная величина, такая как электрический ток, или интегральная, такая как потребленная мощность за период времени или средняя массовая концентрация частиц в атмосфере. Физическая величина может измеряться непосредственно приборами (прямые измерения) либо вычисляться *вычислительными компонентами* на основании результатов измерения других физических величин (косвенные измерения).

Средство измерения (*Measuring_instrument*) — сущность, производящая непосредственное измерение физической величины. В онтологии эта взаимосвязь отражена отношением `<НепосредственноИзмеряет>` (`<MeasuresDirectly>`).

Компонент измерительной системы (*Component*) — входящее в состав системы техническое устройство, выполняющее одну из функций в зависимости от вида компонента. Компоненты подразделяются на измерительные, вычислительные, связующие, вспомогательные и комплексные. Компонент измерительной системы может быть реализован с помощью средства измерений, поэтому в онтологии предусмотрено отношение `<РеализуетсяЧерез>` (`<ImplementedBy>`). Компоненты входят в состав измерительных каналов, эта взаимосвязь отражена отношением `<ПринадлежитИзмерительномуКаналу>` (`<BelongsToChannel>`).

Измерительный компонент (*Measuring_component*) — это устройство, занимающееся сбором первичных результатов измерения. Примерами служат измерительные приборы, первичные и промежуточные измерительные преобразователи, аналоговые “вычислительные” устройства.

Вычислительный компонент (*Computing_component*) выполняет вычисление результатов прямых, косвенных, совместных или совокупных измерений на основании данных средств измерений или измерительных компонентов.

Связующий компонент (*Connecting_component*) предназначен для передачи результатов измерения от одного компонента измерительной системы к другому. Например, проводная линия связи, радиоканал, телефонная линия связи, оптоволокно, спутниковая линия связи.

Вспомогательный компонент (*Auxiliary_component*) — техническое устройство, не участвующее в процессе измерения, но необходимое для нормального функционирования измерительной системы.

Комплексный компонент (*Complex_component*) — совокупность компонентов, завершающая измерительные преобразования, вычислительные и логические операции, предусмотренные процессом измерений и алгоритмами обработки результатов изме-

рений. В качестве примеров можно назвать контроллеры и программно-технические комплексы.

Измерительный канал (*Measuring_channel*) — конструктивно или функционально выделяемая часть измерительной системы, выполняющая весь комплекс преобразований по измерению физической величины. Измерительный канал является совокупностью компонентов, эта взаимосвязь отражена отношением <ПринадлежитИзмерительномуКаналу> (<BelongsToChannel>), определенным для компонента измерительного канала. Кроме того, измерительный канал измеряет некоторую физическую величину, в онтологии это отражено через отношение <Измеряет> (<Measures>). Принадлежность канала измерительной системе отражена отношением <ПринадлежитСистеме> (<BelongsToSystem>).

Измерительная система (*Measuring_system*) — совокупность компонентов, объединяемых в измерительные каналы, предназначенная для измерения набора изменяющихся во времени и распределенных в пространстве физических величин, регистрации и обработки результатов измерений.

Дополнительно в рассматриваемой онтологии введены следующие предположения.

- Отношение <РеализуетсяЧерез> между средствами измерений и компонентами измерительных каналов имеет тип “многие ко многим”. Это означает, что один компонент может быть реализован несколькими средствами измерений (например, компонент, измеряющий трехфазный ток, реализован через три однофазных измерительных трансформатора). В то же время одно средство может реализовывать несколько компонентов (например, если оно подключено к нескольким измерительным каналам).

- Предполагается, что измерительные каналы и их компоненты являются сущностями сугубо логическими, тогда как средства измерений — модели физических сущностей. Это позволяет отделить понятие физического средства как материального объекта от функциональной роли, которую оно играет в процессе измерения.

- *Измерительная система* обладает всеми признаками средств измерения и является их разновидностью (в онтологии этот факт отражен при помощи отношения <is-a>).

- Отношение <Измеряет> между измерительным каналом и физической величиной подчиняется правилу логического вывода: если измерительный канал <Измеряет> физическую величину, то существует принадлежащий ему компонент, <РеализованныйЧерез> средство измерения, которое <НепосредственноИзмеряет> эту физическую величину.

3. Примеры применения онтологии измерительных систем

Разработанная онтология может быть применена в любой предметной области, в которой возникает потребность в автоматизированных процессах измерения. Рассмотрим два примера: система учета электроэнергии и атлас “Атмосферные аэрозоли Сибири”.

Процесс учета энергопотребления на промышленных предприятиях традиционно является предметом метрологического нормирования [4]. Благодаря этому существует устоявшаяся практика применения терминов, введенных в предыдущем разделе, при построении систем учета электроэнергии. *Измерительными компонентами* здесь являются измерительные трансформаторы тока и напряжения и счетчики электроэнергии. *Связующими компонентами* служат каналы передачи данных и телекоммуника-

ционное оборудование. Завершает измерительные каналы вычислительный комплекс (*вычислительный компонент*), который производит расчет количества электроэнергии, потребленной объектом, на основе данных измерений в соответствии с заданными алгоритмами. Основной *физической величиной*, измеряемой системой, является значение расхода электрической энергии за период времени. Онтология, рассматриваемая в настоящей статье, с успехом применялась при разработке крупномасштабной распределенной промышленной системы сбора и учета данных энергопотребления.

Атлас “Атмосферные аэрозоли Сибири” [5] разрабатывается Институтом вычислительных технологий и Институтом химической кинетики и горения СО РАН. Он представляет собой web-портал, аккумулирующий накопленный до настоящего времени эмпирический материал о составе и распределении атмосферных аэрозолей на территории Западной и Восточной Сибири, Алтайского и Красноярского краев. Портал предоставляет возможность коллективной работы с архивами данных и возможность обработки данных специализированными алгоритмами для решения исследовательских задач. К физическим характеристикам атмосферного аэрозоля, измеряемым в атласе, относятся счетная и массовая концентрация аэрозольных частиц, спектр размеров частиц и полное аэрозольное светорассеяние.

Экспериментальный характер исследований в области физики атмосферы позволяет обойтись без долгосрочных трудоемких процедур метрологической аттестации процессов измерения этих характеристик. Поэтому необходимо выполнить явную привязку метрологических терминов к практике формирования эмпирического материала по атмосферным аэрозолям, основываясь на ее физическом содержании. При наличии онтологии измерительных систем такая привязка составляет основной объем информационного моделирования предметной области, поскольку ее результат пригоден к непосредственному применению для проектирования информационно-измерительной системы. Привязка начинается с определения физических величин, измеряемых в контексте атласа “Атмосферные аэрозоли Сибири”, это концентрации частиц в воздухе и другие свойства атмосферных аэрозолей. Далее, измерительными каналами являются станции наблюдения, по которым проводился атмосферный мониторинг, средствами измерения — установленные на них измерительные приборы.

Кроме того, в рамках совместных исследований с Сибирским отделением РАН, Институт космических исследований РАН (г. Москва) проводит сбор данных со станций приема космической информации с помощью низкоорбитальных спутников серии NOAA и Метеор, представляющих спутниковые снимки различных участков Земли. Полученная информация используется в атласе для расчета характеристики отражения света от земной поверхности — альбедо. Здесь измерительными компонентами являются сами спутники, предоставляющие первичные данные измерений. Модуль расчета альбедо по данным снимков, расположенный на сервере атласа, является вычислительным компонентом. Так получается измерительный канал, измеряющий физическую величину альбедо.

Атлас “Атмосферные аэрозоли Сибири” содержит также ряд дополнительных вычислительных компонентов, выполняющих функции подсчета наиболее значимых статистических характеристик химического состава аэрозолей; корреляционного анализа, показывающего взаимосвязь химических элементов и их влияние друг на друга; факторного анализа временных рядов, основанного на корреляционном анализе и позволяющего выделить ряд факторов, которые обуславливают присутствие тех или иных соединений в составе аэрозоля. Результаты расчетов хранятся в единой базе данных ат-

ласа совместно с результатами измерений, с автоматической поддержкой необходимых ссылочных связей и метаданных [6].

4. Варианты использования онтологии измерительных систем

Предложенная онтология может быть использована при разработке любой системы, производящей, обеспечивающей и поддерживающей процессы измерения физических величин. Разработчикам программного обеспечения она служит основой для формирования разнообразных информационных структур, возникающих в ходе технологического процесса проектирования. Рассмотрим несколько примеров.

Словарь предметной области. Онтология содержит общую терминологическую базу предметной области, поэтому разработчики программного обеспечения могут использовать термины из онтологии для документирования своего продукта и формирования пользовательского интерфейса, в том числе и многоязычного. Поэтому предложенная в данной статье онтология измерительных систем поддерживает русский и английский язык.

Отображение на базу данных. Онтология предоставляет набор базовых терминов предметной области, с которыми приходится иметь дело в любом процессе измерения. Поэтому онтология является удобным базисом для разработки схемы данных измерительной системы. Она не является полной, поскольку любой конкретный процесс измерения имеет частные особенности, не задаваемые на уровне онтологии. Тем не менее она определяет базовые понятия, которые в той или иной форме присутствуют или должны присутствовать в любой схеме данных.

Форматы хранения метаданных. Свойства онтологических терминов определяют состав и формат представления метаданных, содержащихся в системе. Эффективная поддержка метаданных является одной из ключевых задач инженерии информационных систем [7]. Привлечение онтологии позволяет повысить эффективность реализации различных средств обработки данных благодаря формированию богатых массивов метайнформации в машинночитаемой форме.

Форматы обмена данными. Открытые форматы обмена данными с внешними системами, основанные на онтологии, существенно упрощают задачу интеграции систем, относящихся к различным областям либо созданных различными разработчиками.

Заключение

При разработке программных систем для автоматизации процессов измерения и сбора данных неизбежно возникает проблема представления метаданных результатов измерений: пространственно-временной привязки, условий проведения измерений, характеристик измерительных приборов. Использование современного подхода к хранению и представлению таких метаданных может значительно упростить процесс разработки программных систем, повысить их качество, упростить процесс интеграции систем из смежных областей. Семантически богатые метаданные в измерительных системах представляют собой стартовую площадку для разработки сложных алгоритмов и интеллектуальных программных агентов.

Онтология измерительных систем может использоваться инженерами по программным системам и разработчиками программного обеспечения как для получения общих знаний о процессе сбора предметных данных, так и для решения ряда практических задач, начиная от проектирования базы данных и заканчивая формированием пользовательского интерфейса системы.

Приведенная в настоящей статье онтология рассчитана на дальнейшее масштабирование и может быть детализирована до любого необходимого уровня. При этом базовые термины измерительных процессов не требуют модификации и не будут затронуты. Это позволит строить производные онтологии, адаптированные под частные предметные области и процессы измерения, с сохранением их взаимной совместимости на уровне ключевых метрологических понятий. Такие онтологии требуются, в частности, в различных задачах автоматизации экологических исследований.

Список литературы

- [1] BACHER R., LEAL D., SCHRODER A. ScadaOnWeb — modelling and web-exchange of process and engineering information // Proc. Intern. Semantic Web Conf. Cagliari, Italy, 2002.
- [2] OWL Web Ontology Language Guide. W3C Working Draft. W3 Consortium, 2003.
<http://www.w3.org/TR/2003/WD-owl-guide-20030331/>
- [3] ГОСТ Р 8.596-2002. Метрологическое обеспечение измерительных систем. Основные положения. М.: Госстандарт России, 2002.
- [4] Димов Ю.В. Метрология, стандартизация и сертификация. СПб.: Питер, 2004.
- [5] Гордов Е.П., Ковалев С.П., Молородов Ю.И., Федотов А.М. Web-система управления знаниями об окружающей среде // Вычисл. технологии. Т. 10. Спецвыпуск. Ч. 2: Тр. Междунар. конф. и школы молодых ученых “Вычислительные и информационные технологии для наук об окружающей среде” (CITES-2005). Новосибирск, 13–15 марта 2005 г. С. 12–19.
- [6] KOVALYOV S.P. Architecture of distributed information-computing system for exploring atmospheric aerosol // Proc. of SPIE. 2005. Vol. 6160. Pt I: Intern. Conf. “Molecular Spectroscopy and Atmospheric Radiative Processes”, Tomsk, June 27–30, 2005. P. 21–26.
- [7] SHANKARANARAYANAN G., EVEN A. The metadata enigma // Comm. ACM. 2006. Vol. 49(2). P. 88–94.

Поступила в редакцию 11 мая 2007 г.