

УДК 519.7

Ю.С. НОВИКОВ

СТРУКТУРИЗАЦИЯ СОСТАВНЫХ ОБЪЕКТОВ ПРИ ФОРМИРОВАНИИ ПРОЦЕССНОГО ПРЕДСТАВЛЕНИЯ ЗНАНИЙ

Предлагается подход к выполнению структуризации составных объектов как элементов процессного представления знаний. При структуризации выполняется последовательное сравнение простых атрибутов объектов и их совокупности на основе их типов и значений, сравнения уникальных атрибутов на основе специализированных алгоритмов, сравнения вложенных объектов.

1. Актуальность

Процессное представление знаний формируется на основе анализа наборов последовательностей событий, имеющих временные метки, и выявления на основе такого анализа структурированных объектов и последовательностей действий, составляющих процесс, а также причинно-следственных зависимостей между такими действиями. Для решения указанных задач анализа традиционно используются методы data mining и, в последние годы, методы бурно развивающегося направления process mining (интеллектуального анализа процессов). Исходными данными при решении таких задач являются наборы последовательностей событий, фиксирующих поведение некоторых объектов и связанных с ними процессов. Интеллектуальный анализ процессов направлен на выявление последовательности взаимодействия таких объектов с учетом скрытых причинно-следственных связей между ними. Результатом данного анализа является процессное представление знаний, отражающее не просто возможные наборы действий в предметной области, но и зависимости между ними, что создает возможность организации логического вывода на процессных знаниях. В то же время выявленные причинно-следственные связи между действиями процесса отражают лишь «внешнюю» природу процессных знаний. Данные зависимости в значительной степени зависят от природы и структуры объектов, с которыми оперирует процесс. Следовательно, структура таких объектов отражает «внутренние» закономернос-

ти процессных знаний, которые в значительной степени ограничивают возможные наборы последовательностей действий процесса. В то же время задача выявления структуры таких объектов, а также определения их влияния на зависимости между действиями процесса до настоящего времени не рассматривалась. Поэтому тема данной работы является актуальной.

2. Анализ последних научных публикаций

Вопросы интеллектуального анализа процессов рассматривались в работах [1,2], где основное внимание уделено обнаружению последовательности действий процесса, а также правил, задающих взаимосвязи между ними. В [3] рассмотрены вопросы выявления структуры сложных наборов данных в базах данных.

Таким образом, вопросы выделения зависимостей и правил при процессном представлении знаний, а также структуризации объектов при проведении интеллектуального анализа данных в литературе рассматриваются разрозненно. В то же время иерархически устроенные сложные объекты являются существенным элементом процессного представления знаний, влияющим на последовательность действий процесса.

3. Постановка задачи исследования

Задача данного исследования заключается в разработке подхода к структуризации составных объектов при формировании процессного представления знаний.

Разработка данного подхода требует: выделения иерархии структурных составляющих составных объектов как элементов процессного представления знаний; определения класса методов (подходов), пригодного для сравнения указанных составляющих и выделения их атрибутов.

Результатом структуризации сложных составных объектов является набор шаблонов и правил, отражающих зависимости между его элементами и непосредственно влияющих на структуру процессного представления знаний.

4. Структуризация составных объектов

Сложные составные объекты являются элементами процессного представления знаний, основаны на структурированных данных и содержат иерархию связанных простых объектов, взаимосвязи между которыми влияют на последовательность процессных действий. Следовательно, возможности манипулирования семантически связанными объектами, которые составляют иерархическую структуру, зависят от наборов их атрибутов, возможных шаблонов и правил их взаимодействия. Примерами предметных областей, для которых характерно манипулирование такими объектами, являются медицина, биология, реинжиниринг бизнес-процессов.

Работа со сложными объектами и построение их иерархии широко используется в области объектно-ориентированного программирования. Ключевыми элементами описания сложных структурированных объектов при этом являются: идентификаторы объектов, их атрибуты, классы, а также механизм наследования. При работе с процессными знаниями меняется наполнение указанных элементов.

Обработка таких структурированных объектов основана на выполнении операций с их атрибутами, например сравнение атрибутов для соответствующих элементов объектов. Такое сравнение в значительной степени зависит от типов сравниваемых атрибутов. Так, для атрибутов с простыми типами первоначально выполняется сравнение типов, а затем сравнение значений с использованием некоторой метрики. Аналогично выполняется сравнение группы атрибутов. В то же время при сравнении специфических типов данный подход слабо применим. В этом случае используются разработанные пользователем специализированные функции или общие принципы сравнения. Дополнительные проблемы возникают при сравнении атрибутов, отличающихся размерами, смысловым наполнением и т.п. В данном случае необходимы компромиссные решения, которые учитывают указанные составляющие. Сравнение иерархии объектов объединяет все указанные типы сравнений атрибутов (рисунок). Сложность таких сравнений зависит от глубины иерархии объектов. Это требует проведения структуризации таких объектов путем выделения соответствующих шаблонов, иерархий и правил взаимодействия их составляющих на основе выявления различий в структуре объектов.



Сравнение сложных составных объектов и их составляющих

Предлагаемый подход к структуризации сложных иерархических объектов заключается в последовательном определении различий между их элементами согласно схеме, представленной на рисунке. Иерархия функций сравнения согласно предлагаемому подходу выглядит следующим образом. На нижнем уровне выполняется сравнение отдельных атрибутов с простыми типами, что позволяет оценить наличие базовых типов в структуре объекта. Для сравнения совокупности всех простых атрибутов целесообразно использовать обобщенную функцию, которая выделяет отличия для значений простых типов объекта в целом. Следующий уровень сравнения – выявление допустимых отклонений, полученных в результате сравнения простых атрибутов иерархического устроенного объекта. Далее необходимо сравнить объекты, составляющие сложный иерархический объект, на основе сравнения их атрибутов, как это было показано ранее. Завершающий этап предлагаемого подхода заключается в попарном сравнении иерархически устроенных объектов на основе выявления различий в выполненной на предыдущих этапах структуризации.

5. Выводы

Предложен подход к структуризации составных объектов как элементов процессного представления знаний. Он предусматривает выполнение структуризации путем последовательного сравнения простых атрибутов объекта и их совокупности на основе их типов и значений, сравнения уникальных атрибутов на основе специализированных алгоритмов, сравнения вложенных объектов. Предлагаемый подход создает возможность выделения структурных элементов сложных составных объектов методами интеллектуального анализа данных и процессов и тем самым обеспечивает условия для формирования процессного представления знаний.

Список литературы: 1. *Process Mining : Discovery, Conformance and Enhancement of Business Processes* / W. M. P. van der Aalst . N.–Y. : Springer Verlag, 2011. 370p. 2. *Aalst, W.M.P. van der Process mining in web services : the websphere case* / Aalst, W.M.P. van der, Verbeek // *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*. 2008. №31(3). P. 45-48. 3. *Han J., Nishio S., and Kawano H. Knowledge discovery in object-oriented and active databases*. In F. Fuchi and T. Yokoi, editors, *Knowledge Building and Knowledge Sharing*, pages 221-230. Ohmsha, Ltd. and IOS Press, 1994. 5.

Поступила в редколлегию 16.11.2012

Новиков Юрий Сергеевич, аспирант кафедры ПИС ХНУРЭ. Адрес: Украина, 61166, Харьков, пр. Ленина, 14, тел. 70-21-446.