
УДК 519.7

С.Ф. ЧАЛЬИЙ, И.Б. БУЦУКИНА

ПРЕДСТАВЛЕНИЕ ЗНАНИЙ В ФОРМЕ ПРОЦЕССОВ С ОГРАНИЧЕНИЯМИ В ЗАДАЧАХ ИНТЕЛЛЕКТУАЛЬНОГО АНАЛИЗА ПРОЦЕССОВ

Предлагается модель представления знаний в форме процессов как развитие скриптового представления знаний. В отличие от существующих, модель отображает знания в виде графа, отражающего возможные состояния процесса и переходы между этими состояниями, а также включает ограничения в форме правил, задающие допустимые последовательности переходов. Совокупность правил задает набор ограничений на выполнение логического вывода на процессных знаниях.

Введение

Интеллектуальный анализ процессов (process mining) направлен на построение их моделей путем анализа последовательностей событий, которые фиксируются информационными системами, и, на этой основе, выделения знаний, отражающих цепочку причинно-следственных зависимостей между событиями и действиями процесса.

Процесс в данном случае можно определить как взаимосвязанный набор последовательностей действий, предназначенных для приема, передачи и обработки материальных, информационных и иных объектов.

В практическом аспекте в качестве процессов как последовательностей действий обычно рассматриваются бизнес-процессы и другие процессы обработки ресурсов, процессы в социальных сетях, процессы, происходящие при функционировании программного обеспечения, процессы мышления, принятия решений т.п.

Результатом интеллектуального анализа процессов (ИАП) обычно является усовершенствование процессов из рассмотренных предметных областей [1,2].

В области ИАП получили дальнейшее развитие подходы к выявлению знаний, которые используются для интеллектуального анализа данных. Однако в целом задачи, решаемые в данной области, не ограничиваются только выделением знаний о процессах и построением на этой основе их моделей «как есть», отражающих их фактическое поведение [2].

Формирование модели процесса «как есть» дает возможность проверить соответствие между последовательностью действий в модели процесса «как должно быть», построенной на этапе проектирования, и фактически выполненной последовательностью действий с учетом существующих внешних ограничений. Проверка соответствия позволяет выявить недопустимые отклонения. Все это требует построения модели представления знаний о

процессах. В то же время существующие модели представления знаний – от продукционной до скриптовой [3,4] не учитывают влияние внешних ограничений при выполнении логического вывода, что и свидетельствует об актуальности проблематики данной работы.

Постановка задачи

В настоящее время выделяют три основных задачи, которые решаются с использованием методов интеллектуального анализа процессов:

1) выявление знаний о зависимостях между действиями процесса и построение модели фактически выполняющегося процесса в виде графа, отражающего возможные последовательности его действий;

2) проверка соответствия модели процесса, построенной на этапе его проектирования, и фактически выполняющегося процесса на основе знаний о допустимых последовательностях действий процесса;

3) локальное усовершенствование процесса или же его реинжиниринг на основе использования его первоначальной модели «как должно быть» и результатов решения первой и второй задач.

При решении указанных задач известными методами возникает ряд трудностей, связанных с многовариантностью представления процессов. Последняя определяется использованием правил, задающих изменение последовательности действий во время выполнения процесса. Совокупность таких правил совместно с последовательностью действий процесса и задает базу знаний об одном процессе.

Отметим, что эти правила выходят за пределы логики процесса, определяющей его обычную последовательность действий. Они фактически задают систему ограничений при поведении процесса, что и приводит нас к необходимости представления процессов с использованием совокупности правил как системы ограничений. Решение данной задачи позволит выявлять знания о процессах на основе логов событий с использованием методов программирования в ограничениях (constraint-based programming).

Алгебраическая модель представления процессных знаний

При формализации поведения систем параллельной работы и распределенных систем в области компьютерных наук и искусственного интеллекта используется теория процессов, в основе которой лежит математический аппарат алгебры процессов. Собственно термин «процесс» описывает поведение моделируемой системы (будь то информационная, техническая, биологическая или человеческая) с возможностью распараллеливания.

В рамках теории процессов выполняется их моделирование и верификация. При моделировании формализуется представление процессов. При верификации осуществляется проверка процесса, например на предмет соответствия реального поведения процесса и имеющейся модели. Для выполнения верификации необходимо определить критерии сопоставления процессов, которые базируются на семантике соответствующей процессной алгебры.

Традиционно при моделировании знаний о процессах основное внимание уделяется рассмотрению последовательности действий. Конкретный смысл этих действий не отражается в модели. Другими словами, для модели при таком подходе несущественно, отражает ли процесс последовательность действий программы, последовательность работ бизнес-процесса или же набор действий пользователя при общении в социальной сети.

Таким образом, при описании функционирования распределенных динамических систем в виде системы процессов каждый из таких процессов может быть представлен в виде набора действий, которые выполняет система, а также порядка выполнения этих действий и, возможно, дополнительных атрибутов действий. Выполнение процесса представляет собой последовательность переходов между состояниями. Каждый переход отражает изменение состояния процесса в результате выполнения его действий в дискретные моменты времени. Следовательно, рассматриваемые процессы являются дискретными. Внутренние особенности действий также не являются важными, поскольку их результат моделируется в виде изменения состояний процесса.

Детализуя процесс как набор возможных последовательностей действий, которые имеют входы и выходы, модель представления процессных знаний объединяет статическую и динамическую составляющие.

Первая из них включает в себя: входы как набор (подмножеств) переменных; выходы как набор переменных; начальное состояние как множество переменных, отражающих состояние объектов и действий процесса; набор допустимых конечных состояний (либо набор ограничений на допустимые конечные состояния); последовательность переходов между состояниями, которая определяется действиями и входами процесса; порядок взаимодействия с другими процессами, который представляется в форме правил-ограничений; порядок организации параллельного и (или) распределенного выполнения процесса (или нескольких процессов), который формализуется в виде правил взаимодействия.

Статическая составляющая процесса описывается графом, вершины которого отражают состояния моделируемой динамической системы, дуги – переходы между этими состояниями. Переходы связаны с выполнением действий над объектами процесса.

Динамическая составляющая знаний о процессе определяет его функционирование и представляет собой последовательность переходов между его состояниями, причем каждый из таких переходов связан с выполнением соответствующих действий. Первоначально процесс находится в начальном состоянии s_0 , в дальнейшем на каждом i - шаге – в состоянии s_i . В дальнейшем переход между состояниями происходит следующим образом. В i - состоянии процесс выбирает дугу с действием a_i , которое может быть выполнено в данный момент времени. Выбор конкретного действия задается описанными ранее правилами.

Затем действие a_i выполняется и процесс переходит в следующее состояние s_{i+1} . Процесс продолжается до достижения одного из конечных состояний. В том случае, если из состояния s_i существует хотя бы одна дуга в последующую вершину s_{i+1} графа процесса, но отсутствует готовое к выполнению действие, то процесс приостанавливается и ожидает готовности действия a_i . При его готовности процесс возобновляет выполнение.

Выполнение одного из возможных в текущем состоянии действий процесса зависит от его входного состояния и набора выполненных действий, а также имеющихся ограничений (правил). Конечным состоянием является такое, из которого отсутствуют дуги в другие состояния процесса.

Реализация каждого процесса (трасса процесса) представляет собой последовательность действий, переводящих процесс из начального состояния в конечное. В общем случае каждый процесс обладает множеством трасс.

При выполнении процесса могут быть достигнуты не все состояния. Состояние s_i будет достижимым для процесса P в том случае, если существует конечная непустая последовательность переходов из начального состояния s_0 в состояние s_i . Очевидно, что совокупность достижимых состояний процесса и соединяющих их дуг представляет собой достижимую часть процесс при заданных ограничениях на его выполнение.

Совокупность всех состояний процесса составляет пространство процессов. При этом любое состояние в данном пространстве потенциально может быть начальным состоянием процесса.

Для каждого процесса P могут быть заданы ограничения с учетом атрибутов операций. Так, ограничение по объектам определяется следующим образом. Пусть P – процесс, а O^* – подмножество объектов, над которым выполняются операции A данного процесса.

Тогда ограничением процесса P по множеству объектов O^* является граф P^* , который получен из графа P удалением тех операций, которые связаны с обработкой объектов из подмножества O^* .

Подведя итог рассмотренным особенностям формализации процессных знаний, можно сделать вывод о том, что формальная модель процессных знаний P отражает сочетание декларативного и процедурного подходов к представлению знаний. Действительно, процедурные знания определяют последовательность действий по достижению искомой цели. Декларативные же знания охватывают совокупность правил, задающих локальные и глобальные ограничения на последовательность действий процесса. Следовательно, проце-

дурные знания задают цепочки действий процесса, а декларативные – условия запуска и порядок применения указанных цепочек действий.

Таким образом, приведенная формализация дает возможность сформировать процессное представление знаний, которое характеризуется следующими особенностями:

- каждое действие процесса, представленное в виде <состояние - действие-состояние>, полностью соответствует традиционному правилу продукции для представления знаний;

- представление процессов в виде взаимосвязанных наборов действий задает множество возможных вариантов логического вывода на основе процессных знаний и соответствует процедурному представлению знаний;

- структуризация отдельных действий средствами алгебры частичного порядка позволяет учесть их атрибуты при организации логического вывода на процессах, что дает возможность задать локальные ограничения на допустимые пути логического вывода на процессах;

- представление последовательности действий процесса в виде набора связанных причинно-следственными связями правил дает возможность задать глобальные ограничения на поведение процесса в виде таких же правил, что позволяет динамически изменять вывод на процессных знаниях, накладывая ограничения во время реализации процесса, например при изменении целей, внешних условий и т.п.

Отметим, что выделенные особенности процессной модели как модели, основанной на цепочке правил, позволяет использовать методы *constraint-based programming* при формализации процесса на основе анализа логов событий. Основная идея программирования с ограничениями состоит в решении комбинаторных задач путем нахождения такой комбинации ограничений, которая удовлетворяет набору заранее заданных ограничений. Иными словами, модель задачи задается в терминах ограничений. Общий алгоритм решения определяет последовательность действий по решению задачи.

Выводы

Предложена модель представления знаний в форме процессов, которая является дальнейшим развитием скриптового представления знаний. Она отличается наличием ограничений в форме правил, что обеспечивает сочетание преимуществ декларативного и процедурного подходов к формализации знаний. В соответствии с предложенной моделью знания представляются в виде графа, отражающего возможные состояния процесса и переходы между ними. Каждый такой переход может быть при необходимости представлен в виде логического правила. Совокупность правил задает набор ограничений на выполнение логического вывода на процессных знаниях. Это дает возможность решать задачи интеллектуального анализа процессов методами программирования с ограничениями, где ключевым является задание собственно системы ограничений.

Список литературы: 1. *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.* 2. *W. M. P. van der Aalst. Process Mining : Discovery, Conformance and Enhancement of Business Processes / W. M. P. van der Aalst. N.–Y. : Springer Verlag. 2011. 370 p.* 3. *Представление и использование знаний: Пер. с япон./Под ред. Х. Уэно, М. Исидзука. М.: Мир, 1989. 220 с.* 4. *Sowa J.F. Knowledge Representation: Logical, Philosophical, and Computational Foundations / Sowa J.F. Brooks-Cole, 2000.*