

НАУКОЕМКИЕ ИССЛЕДОВАНИЯ И РАЗРАБОТКИ

НЕКОТОРЫЕ ПОДХОДЫ К АНАЛИЗУ И АВТОМАТИЗАЦИИ ПРОЕКТИРОВАНИЯ ПОТОКОВ РАБОТ

М. А. Филатов

Московский инженерно-физический институт, Москва

Рассмотрена математическая модель потока работ, основанная на использовании аппарата сетей Петри. Предлагается использование концепции наследования для решения задач анализа потоков работ и сохранения семантики при проектировании описаний бизнес-процессов. Описывается разработанное автором CASE-средство ER&LC 2.0, интегрированное с ERP-системой MBS Ахарт и используемое для автоматизации проектирования бизнес-процессов, создания и доработки ИС.

Технологии потоков работ (workflow technology) получили в последнее десятилетие широкое распространение. Системы управления потоками работ (WfMS — Workflow Management systems) используются во многих прикладных областях, таких как автоматизация офисной деятельности, управление производственными процессами и т. д. Многие современные системы управления ресурсами предприятия класса ERP (Enterprise Resource Planning) содержат модуль управления потоками работ. В целом можно наблюдать изменение парадигмы современных информационных систем (ИС) от автоматизации отдельных задач к комплексной автоматизации бизнес-процессов. Одним из магистральных направлений развития архитектуры современных ИС является использование веб-сервисов (web-services) — слабосвязанных программных компонентов, взаимодействующих между собой с помощью протоколов на основе XML (eXtensible Markup Language). ИС, построенная на основе веб-сервисов, будет представлять собой набор распределенных взаимодействующих потоков работ.

Таким образом, технология потоков работ представляет собой молодую, быстро развивающуюся область информационных технологий, что определяет высокую актуальность исследований вопросов анализа и автоматизации проектирования потоков работ.

Основные определения и стандарты

В 1993 г. была организована Workflow Management Coalition (WfMC) — некоммерче-

ская организация, цель которой — расширение возможностей технологий потоков работ путем разработки единых стандартов.

Часть стандартов WfMC являются открытыми, некоторые другие доступны только членам консорциума. Основное достижение WfMC — создание словаря Terminology and Glossary [1], где определяются основные термины, относящиеся к workflow-системам, и который является на сегодняшний день стандартом де-факто.

Бизнес-процесс (Business Process) — совокупность одной или более связанных между собой процедур или функций, которые совместно реализуют некоторую задачу бизнеса или политическую цель, обычно в рамках организационной структуры, описывающей функциональные роли и отношения [1].

Поток работ (Workflow) — автоматизированный полностью или частично бизнес-процесс, в котором документы, информация или задания передаются от одного участника другому в соответствии с набором процедурных правил [1].

Описание бизнес-процесса (Process definition) — представление бизнес-процесса в форме, поддерживающей автоматизированное манипулирование (например моделирование) или реализацию системой управления потоками работ (СУПР). Описание бизнес-процесса состоит из связанных в некоторую последовательность функций и их отношений, критериев начала и завершения процесса, а также информации об отдельных функциях, например, об участниках, данных и т. д. [1].

В стандарте Workflow Reference Model [2] описывается следующая общая архитектура СУПР:

- ядро системы, содержащее набор выполняемых экземпляров бизнес-процессов;
- редактор описаний бизнес-процессов;
- клиентское приложение, с помощью которого ядро взаимодействует с пользователями;
- внешние приложения, вызываемые системой;
- административное приложение.

Все компоненты взаимодействуют напрямую только с ядром системы, которое выполняет роль посредника между компонентами. Стандарт не касается вопроса, как должны быть устроены компоненты. В нем описываются интерфейсы взаимодействия компонентов с ядром системы.

В 2000 г. консорциум OMG (Object Management Group) выпустил документ Workflow Management Facility Specification [3]. В нем построены основы архитектуры ядра СУПР, на языке IDL определены основные интерфейсы многих компонентов.

В 1999 г. коалицией WfMC был разработан язык определения бизнес-процессов WPDЛ, основанный на нотации Бэкуса—Наура. В 2002 г. язык WPDЛ был переписан, его новая версия (XPDL) использует синтаксис XML.

Языки WPDЛ и XPDL не имеют графической нотации и не основаны на формальной модели, что препятствует решению задач анализа и моделирования описаний бизнес-процессов.

Современные подходы к описанию и анализу бизнес-процессов

В настоящее время наибольшее распространение получили следующие подходы к описанию бизнес-процессов:

- **π -исчисление.** Разработано в конце 80-х годов XX века Робинот Милнерот, основано на алгебре параллельных процессов (одна из модификаций λ -исчисления). В настоящее время π -исчисление является перспективной, хотя и очень молодой и развивающейся теорией, в ней много открытых вопросов и нерешенных проблем. Разработчики таких языков, как BPML и BPEL4WS, утверждают, что так как в их основе лежит солидная математическая теория — π -исчисление, эти языки обладают очень высокой выразительной мощностью. Однако связь данных языков с π -исчислением является скорее маркетинговым ходом.

- **UML.** Значительное развитие получил также подход, связанный с применением UML для описания бизнес-процессов (диаграммы деятельности). Диаграмма деятельности — особая форма конечных автоматов, которая отражает процесс вычисления и потоки работ. Основным недостатком диаграмм деятельности является отсутствие возможности исполнения модели.

- **Сети Петри.** Могут использоваться расширенные сети Петри и методы динамического

функционального анализа на основе расширенных сетей Петри.

Наибольший практический и теоретический интерес связан с использованием сетей Петри, так как они имеют:

формальную семантику. Расширения UML не имеют формальной семантики, что не позволяет проводить анализ создаваемых с их помощью описаний бизнес-процессов;

графическое представление. Любое формальное описание сети Петри однозначно преобразуется в графическую модель, которая может быть использована для обсуждения с конечными пользователями. В то же время π -исчисление не имеет графического представления;

значительный аналитический аппарат. Математический аппарат сетей Петри развивается с 60-х годов XX века, были получены значительные результаты как в теории сетей Петри, так и в использовании сетей Петри в различных прикладных областях. В настоящее время создается международный стандарт ISO/IEC 15909 [4], описывающий моделирование систем с использованием сетей Петри. π -исчисление получило распространение только в начале 90-х годов, многие теоретические вопросы данной концепции еще не требуют исследования.

Использование аппарата сетей Петри для описания и анализа бизнес-процессов

Существующие подходы к представлению потоков работ сетями Петри

Исследования применимости аппарата сетей Петри к представлению потоков работ ведутся с начала 80-х годов [5]. Наибольшую известность приобрели исследования голландского ученого В. Аалста [5, 6], предложившего в середине 90-х годов концепцию WfPN (Workflow Petri Nets). Основные идеи концепции WfPN В. Аалста: каждая задача потока работ преобразуется в переход сети, а состояние — в позицию сети. Сети Петри, представляющие потоки работ, выделяют в отдельный класс WfPN (Workflow Petri Nets), так как они обладают особыми свойствами, обуславливаемыми ограничениями, накладываемыми им на структуру потока работ. WfPN обладает следующими свойствами:

имеет одну начальную и одну конечную позицию;

граф сети Петри является сильно связным.

Базовым свойством, используемым при анализе WfPN, является надежность (soundness), если:

для любой маркировки сети существует последовательность выполнений сети, приводящая к конечной маркировке, содержащей единственную фишку в конечной позиции;

WfPN не содержит висячих переходов.

Доказано, что надежная WfPN (soundness WfPN) не содержит бесконечных циклов и тупиковых ситуаций [6].



WfPN В. Аалста обладают двумя основными недостатками:

так как каждая задача моделируется одним переходом, невозможно более детальное специфицирование жизненного цикла исполнения задачи (начато исполнение, отменена, подтверждена и т. д.);

поток работ может иметь несколько семантически различных начальных и завершающих состояний, поэтому ограничение на одну начальную и конечную позицию является искусственным.

Предлагается вариант WfPN, который сохраняет жизненный цикл исполнения задачи и снимает ограничение на единственность начальной и конечной позиций.

Формализация понятия "поток работ"

Для описания задач потока работ будет использоваться транзакционный подход, в частности алгоритм 2PC (2-phase commit) [7].

Определим поток работ как $W = \langle T, R \rangle$, где $T = \{t_1, t_2, \dots, t_n\}$ — множество задач, каждая задача $t_i = \{op_{i1}, op_{i2}, \dots, op_{ik}\}$ — множество операций. Выделяется подмножество базовых операций $OP_{\text{баз}} = \{\text{начать, подготовить к завершению, отменить, завершить}\}$. Все операции могут быть начаты только после исполнения операции "начать" и должны быть завершены после исполнения операций "отменить" или "завершить". Каждая задача t_i может находиться в одном из состояний $S_{ti} = \{\text{начата, исполняется, выполнена, подтверждена, отменена}\}$. Состояние потока работ $S_w = \{S_{t1}, S_{t2}, \dots, S_{tn}\}$ — множество состояний задач потока работ.

Выделяются следующие типы зависимостей между задачами [7]:

зависимости последовательности выполнения (control flow). Представляют набор условий, определяющих возможность перехода от задачи t_i , находящейся в состоянии S_{ti} , к задаче t_j в состоянии S_{tj} ;

временные зависимости. Выделяют два типа временных зависимостей — абсолютные (данная задача должна начать выполняться в строго определенное время) и относительные (задача t_j должна начать выполняться не более чем через 10 мин после завершения t_i).

Основные понятия сетей Петри

Сеть Петри представляет собой ориентированный граф с вершинами двух типов (позициями и переходами), в котором дуги могут соединять только вершины разных типов (двудольный ориентированный мультиграф). В позиции сети помещаются специальные маркеры, перемещение которых и отображает динамику моделируемой системы. Изменение маркировки (движение маркеров) происходит в результате

выполнения (срабатывания) перехода на основе соответствующего внешнего события. Переход срабатывает, если во всех его входных позициях имеются маркеры и происходит соответствующее переходу событие. При этом из каждой входной позиции срабатывающего перехода маркер удаляется, а в каждую выходную позицию — заносится.

Сеть Петри декомпозирует систему на активные (переходы) и пассивные (позиции — хранилища маркеров) элементы. Формальное определение СП [7]: $СП = \langle P, T, I, O, \mu \rangle$,

где $P = \{p_1, p_2, \dots, p_n\}$ — конечное множество позиций, $n \geq 0$;

$T = \{t_1, t_2, \dots, t_m\}$ — конечное множество переходов, $m \geq 0$;

множество переходов и позиций не пересекаются, $P \cap T = \emptyset$;

I, O — входная и выходная функции, соответственно, являются отображением переходов в комплексы позиций.

Маркировка $\mu = (\mu_1, \dots, \mu_n)$, где $n = |P|$, $\mu = {}^P P \rightarrow N$ — функция, отображающая множество позиций P в множество неотрицательных целых чисел N . Маркировка μ — присвоение фишек позициям СП. Вектор μ определяет для каждой позиции μ СП количество фишек в этой позиции.

Одна из наиболее важных задач анализа сетей Петри — задача достижимости, т. е. определения, является ли маркировка μ' достижимой из μ (в результате исполнения последовательности переходов). Для произвольной сети Петри задача достижимости, хотя и является разрешимой, но соответствующий алгоритм обладает при этом экспоненциальной сложностью.

Используемые расширения сетей Петри

Существует множество расширений сетей Петри, которые были предложены для решения определенных задач, с которыми встретились исследователи при попытках моделировать реальные системы. Акцент в этих расширениях сделан на моделирование, а не на теоретическую мощност.

Основной проблемой, являющейся ограничением мощности моделирования сетями Петри, является нечувствительность условий запуска переходов к числу фишек в позиции. Более конкретно, ограничение на моделирование с помощью сетей Петри состоит в неспособности проверить на наличие точно определенной маркировки в некоторой неограниченной позиции и осуществить действие в зависимости от результатов проверки. Данное ограничение известно как неспособность проверки на нулевую маркировку в некоторой позиции, и поэтому это свойство известно как проверка на нуль. Сети Петри

не могут проверить неограниченную позицию на нуль.

Наиболее известным расширением сетей Петри, которое допускает проверку на нуль, являются сдерживающие дуги. Правило запуска перехода изменяется следующим образом: переход является разрешенным, когда фишки присутствуют во всех его обычных входах и отсутствуют в сдерживающих входах. Переход запускается удалением фишек из всех его обычных входов [8].

Для описания бизнес-процессов наибольшее применение нашли следующие расширения:

временные сети Петри. Важная особенность сетей Петри — это их асинхронная природа. В сети Петри отсутствует понятие времени или измерения времени. Если в какой-то момент времени разрешено более одного перехода, то любой из нескольких разрешенных переходов может быть запущен. Во временных сетях Петри каждому переходу t_j сопоставляются два момента времени τ_{1j} , τ_{2j} . Переход t_j может быть запущен, если он был разрешен к моменту времени τ_{1j} . Если он разрешен, то он должен быть запущен к моменту времени τ_{2j} ;

цветные сети Петри. Расширение обычных сетей в цветные заключается в добавлении следующей информации к компонентам сети: мар-

керы вместо простого обозначения содержимого места преобразуются в объект, который может содержать один или более параметров, каждый из которых может принимать дискретный набор значений. Чтобы отличать маркеры различных типов, их окрашивают в различные цвета;

иерархические сети Петри. При создании описаний сложных бизнес-процессов целесообразно разделять их на отдельные подпроцессы, соответствующие вложенной сети Петри.

WF-сети Петри

Каждая задача t_i потока работ W представляется сетью Петри. Состояния задачи t_i представляются позициями сети, операции — переходами сети (рис. 1).

Для моделирования зависимостей последовательности выполнения между задачами t_i и t_j вводится промежуточное состояние BUF_{ij} (рис. 2). Цель данного состояния — гарантировать, что переход TR_j не будет исполнен до того, как исполнен переход TR_i .

Для моделирования временных зависимостей возможностей классических сетей Петри недостаточно, необходимо использовать временные расширения сетей Петри.

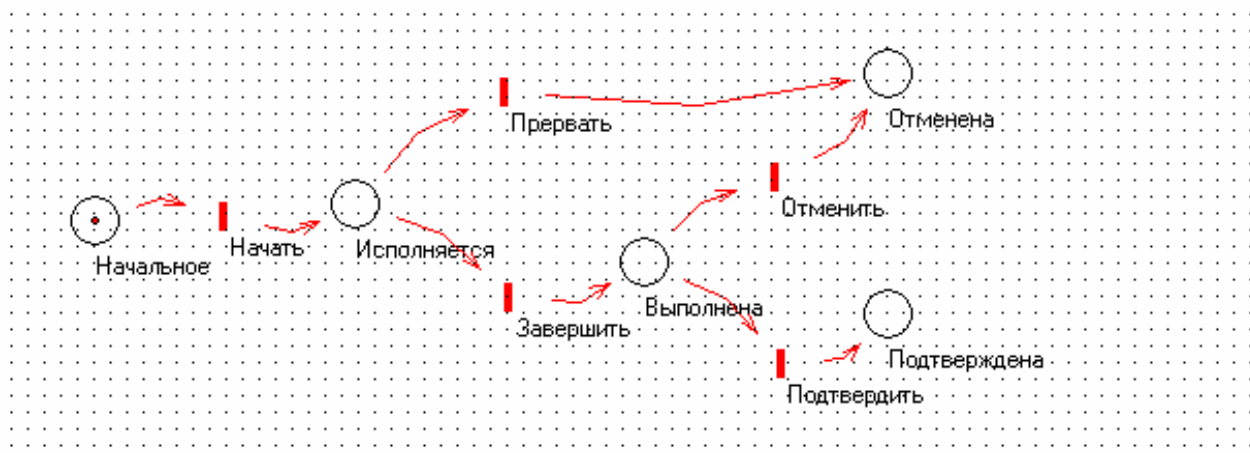


Рис. 1. WF-сеть Петри

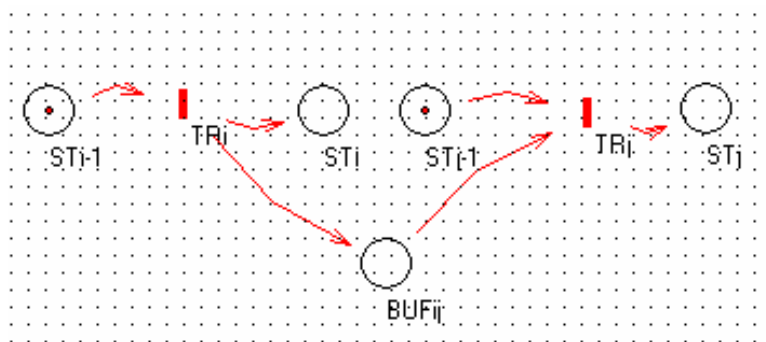


Рис. 2. Зависимость последовательности выполнения между задачами t_i и t_j

Анализ WF-сетей

Рассмотрим одну из основных задачу анализа потоков работ — определение корректного завершения WF-сети. Поток работ корректно завершается, если все его возможные конечные состояния допустимы, т. е. для заданного множества $S'w_{\text{кон}}$ — недопустимые состояния потока работ необходимо проверить, что $S'w_{\text{кон}} \not\subset Sw_{\text{кон}}$, где $Sw_{\text{кон}}$ — множество конечных состояний потока работ. В терминах теории сетей Петри необходимо определить достижимость $S'w_{\text{кон}}$. Для произвольной сети Петри алгоритм проверки имеет экспоненциальную сложность, но можно показать, что описанные выше WF-сети относятся к особому подклассу сетей Петри — ациклическим сетям Петри.

Теорема 1. WF-сеть Петри $C(W)$, спроектированная по правилам "Формализация понятия потока работ", является ациклической. Доказательство не приводится из-за его длины. Для данного подкласса сетей Петри доказана теорема 2.

Теорема 2. В ациклической сети Петри $C = \langle P, T, I, O \rangle$ маркировка μ' достижима из $\mu \Leftrightarrow \exists X = (X_1, X_2, \dots, X_i, \dots, X_{|P|})$, $X_i \in N$ (N — множество натуральных чисел), что является решением уравнения $\mu' = \mu + \mu D$, где D — матрица инцидентий СП. (Доказательство см. в [9].)

Основываясь на теоремах 1 и 2, мы разработали следующий алгоритм проверки WF-сети Петри на корректность завершения.

Алгоритм 1. (Проверка WF-сети Петри на корректность завершения)

Для данной маркировки μ WF-сети Петри $C = \langle P, T, I, O \rangle$ проверить все состояния $s_i \in S'w_{\text{кон}}$ на следующее условие:

для соответствующей s_i маркировки μ_i решить уравнение $\mu_i = \mu + \mu D$;

в решении $X = (X_1, X_2, \dots, X_i, \dots, X_{|P|})$ все X_i — натуральные числа;

если хотя бы для одного состояния данное условие не выполняется, то WF-сеть C завершается некорректно, иначе WF-сеть C завершается корректно.

Концепция наследования потоков работ

Несмотря на широкое практическое использование, ряд важнейших проблем, касающихся систем управления потоками работ, до сих пор не имеет окончательного решения. К этим проблемам относятся [5, 6]:

преобразование с сохранением свойств. Одним из существенных преимуществ применения потоков работ является возможность изменения бизнес-процессов динамически во время выполнения предыдущей версии бизнес-процесса. При этом необходимо избежать возможных аномалий. До сих пор не предложено универсальных методов преобразования одного потока работы в

другой, сохраняющий заданные существенные свойства. Многие современные ERP-системы (SAP, Oracle E_Business Suite и т. д.) имеют так называемые ссылочные модели, содержащие лучшие практики бизнес-процессов, автоматизируемых с использованием данных систем. Тем не менее практически при любом внедрении возникает необходимость той или иной модификации стандартных бизнес-процессов с учетом нужд данного предприятия, при этом необходимо сохранить семантику базового бизнес-процесса;

межорганизационные потоки работ. С развитием электронной коммерции, CALS-технологий все большее количество бизнес-процессов становятся межорганизационными, объединяют бизнес-процессы различных предприятий. Создание и выполнение межорганизационных потоков работ связано с тем, что, с одной стороны, желательно оптимизировать общий поток работ, унифицировать его компоненты, а с другой — необходимо учитывать особенности бизнес-процессов каждой из организаций-участниц.

Одним из наиболее перспективных подходов к решению поставленных выше проблем, возникающих при использовании технологии потоков работ, представляется введение концепции "наследование". Данная концепция получила широкое распространение в теории объектно-ориентированного программирования. Приведем определение наследования, принятое в UML [10]:

Наследование — механизм, с помощью которого более специализированные элементы заимствуют структуру и поведение более общих элементов.

Общее описание называется предком, более детализированное — потомком. Наследование позволяет существенно снизить размер модели и, что более важно, снижает риск возникновения ошибок после внесения в модель изменений.

Поток работ является потомком другого потока работ, если соккрытие или блокировка задач (переходов WfPN-сети) потока-потомка приводит к потоку-предку. Потоки-предки могут быть использованы для определения минимального набора требований, которому должен соответствовать поток-потомок. Основой одного из возможных подходов к решению задачи о том, является ли один поток работ наследником другого, могут служить результаты, полученные в теории языков сетей Петри, которые также служат основой для определения преобразований, сохраняющих наследование [8].

Введенная концепция наследования позволяет предложить некоторые подходы к решению описанных выше проблем:

преобразование с сохранением свойств. Если при преобразовании полученный поток работ является наследником исходного, то он сохраняет свойства исходного потока. Рекомендательный производителем ERP-системы поток работ

является предком для потока, создаваемого при модификации системы при внедрении, что позволяет сохранить семантику стандартного потока работ;

межорганизационные потоки работ. Можно предложить следующую трехшаговую стратегию создания межорганизационного потока работ:

А — создание общедоступного потока работ, описывающего основные характеристики межорганизационного потока работ, согласование данного потока работ всеми организациями,

В — выделение частей общедоступного потока работ, относящихся к каждой организации,

С — создание каждой организацией собственного потока работ, который является наследником соответствующей части общедоступного потока работ.

CASE-средство анализа потоков работ и проектирования ИС ER&LC 2.0

С 1999 г. автор данной статьи участвовал в разработке CASE-средства ER&LC, новая версия которого 2.0 [9, 10] содержит функциональность, позволяющую описывать и верифицировать потоки работ, оценивать эффективность их выполнения с использованием стоимостного анализа, создавать ИС, автоматизирующие выполнение описанных потоков работ.

ER&LC 2.0 состоит из следующих модулей:

- модуль создания модели данных;
- модуль создания жизненных циклов сущностей;
- модуль анализа и моделирования потоков работ;

- модуль генерации ИС;
- модуль реинжиниринга.

В качестве математического аппарата для моделирования потоков работ в CASE-средстве используются цветные иерархические WfPN. Цвета в WfPN используются для интеграции модели данных и модели бизнес-процессов.

На рис. 3 представлен стандартный процесс обработки заказа на отгрузку по предоплате, созданный в CASE-средстве ER&LC 2.0. Сначала создается спецификация, определяющая товар, который заказывает клиент, после этого заказ комплектуется на складе поставщика, а клиент оплачивает счет, после чего происходит отгрузка товара клиенту.

Основной особенностью данного процесса является то, что отгрузка товара произойдет только в том случае, если заказ скомплектован и оплачен, выполнения одного из этих условий недостаточно, так как переход "Отгрузка" выполнится только в том случае, если в позициях C3 и C4 будут маркеры.

Для удобства при моделировании исполнения WfPN используются следующие расширения:

- задание приоритетов выполнения переходов. Алгоритм исполнения сети определяет все допустимые переходы и выбирает для запуска переход с минимальным значением характеристики приоритета;
- сдерживающие дуги;
- возможность установки количества шагов выполнения сети. Используется для автоматического исполнения сети.

Для настройки параметров WfPN используется форма, показанная на рис. 4.

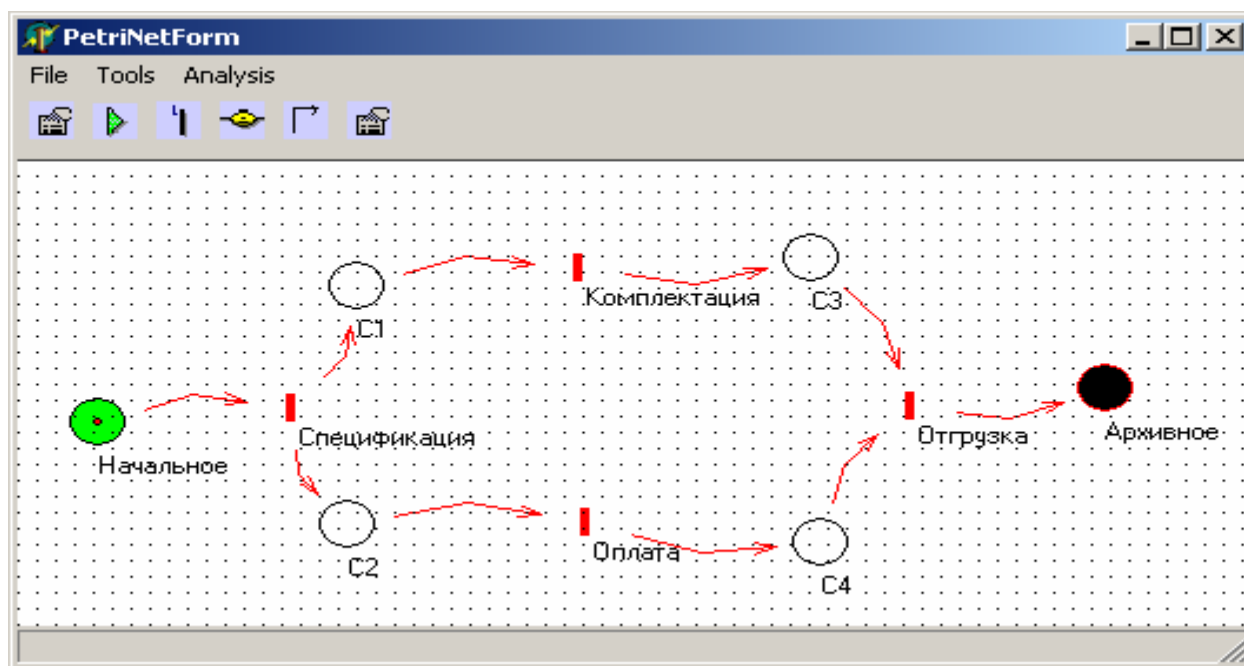


Рис. 3. Процесс обработки заказа по предоплате

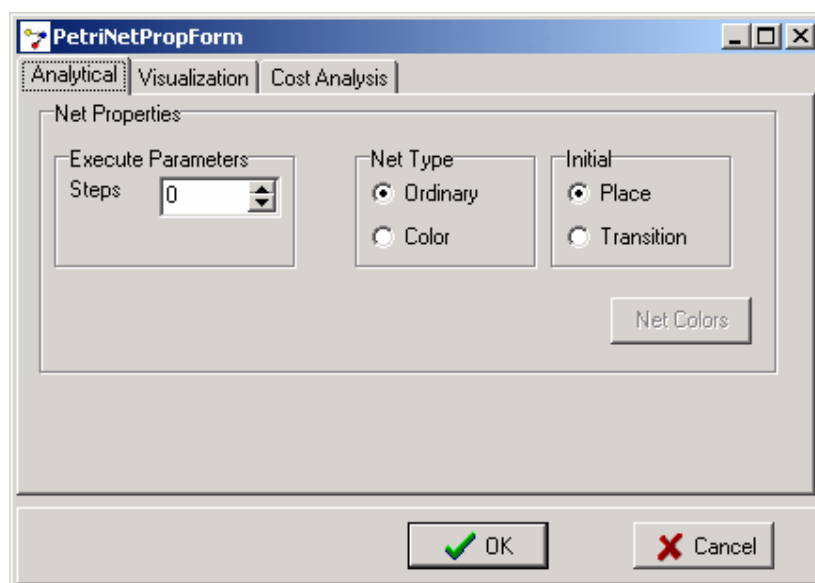


Рис. 4. Форма для настройки параметров сети Петри

Осуществлена интеграция ER&LC 2.0 с ERP-системой MBS Ахарта, что позволяет использовать CASE-систему в процессе внедрения MBS Ахарта на предприятии.

Модуль проектирования потоков работ может быть также использован для моделирования и анализа любых сетей Петри и систем, представляемых сетями Петри. На рис. 5, а показана сеть Петри, созданная с использованием ER&LC 2.0.

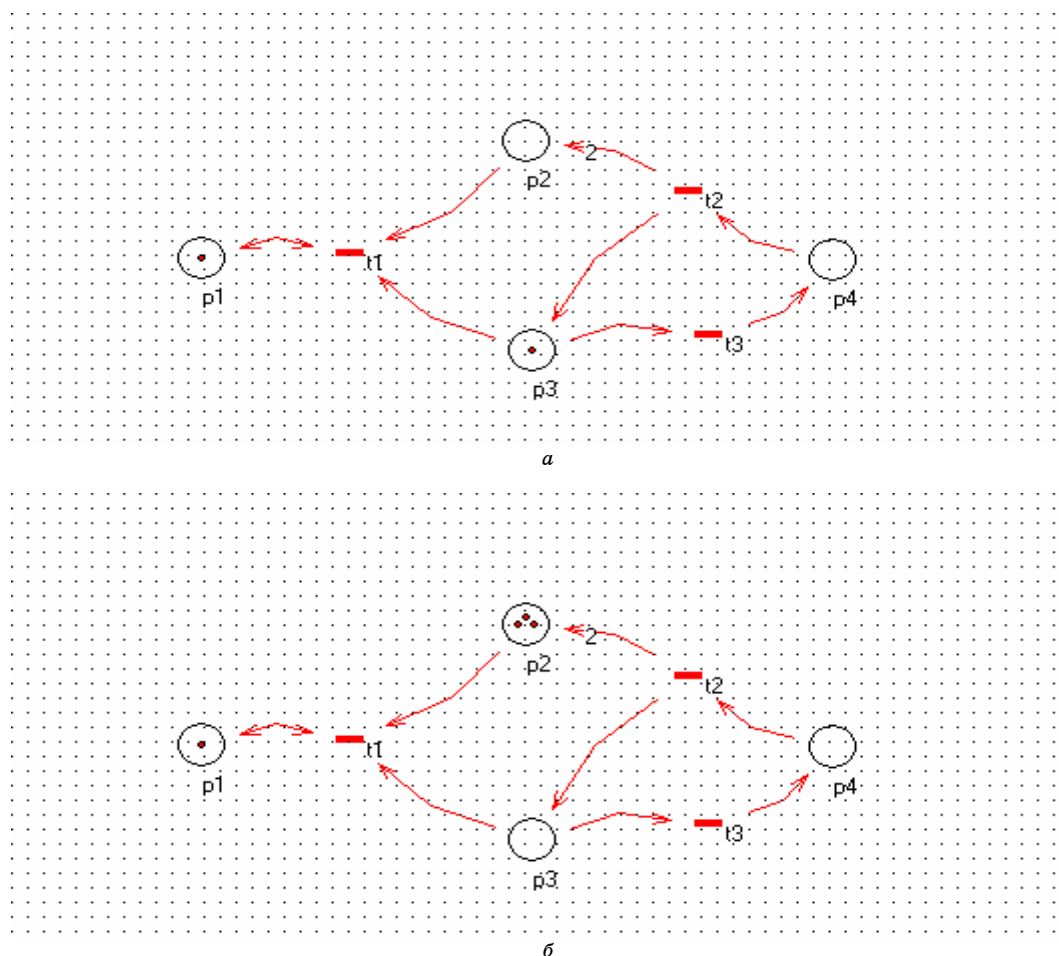


Рис. 5. Состояния сети Петри в ER&LC 2.0:
а — начальное; б — конечное

В начальной маркировке $\mu = (1,0,1,0)$ переход t_3 разрешен и приводит к маркировке $\mu' = (1,0,0,1)$. Последовательность $\sigma = t_3, t_2, t_3, t_2, t_1$ представляется вектором запусков $f(\sigma) = (1,2,2)$, получаем маркировку $\mu_{\text{рез}} = (1,3,0,0)$. Результирующая сеть Петри приведена на рис. 5, б.

Использование аппарата сетей Петри позволяет решить ряд важнейших задач анализа потоков работ. Нами предложен алгоритм определения корректного завершения потока работ. Данный алгоритм использует результаты, полученные в [1, 2, 9].

Использование концепции наследования в дополнение к аппарату сетей Петри, предложенное в данной работе дает дополнительный инструментарий для анализа сложных потоков работ.

Литература

1. WfMC. Terminology and glossary — http://www.wfmc.org/standards/docs/TC-1011_term_glossary_v3.pdf

2. WfMC. Workflow Reference Model — http://www.wfmc.org/standards/docs/TC00-1003_10_1994.pdf
3. OMG. Workflow Management Facility Specification — <http://www.omg.org/docs/formal/00-05-02.pdf>
4. High-Level Petri Nets. Международный стандарт ISO/IEC 15909, версия 4.7.1 — <http://www.informatik.hu-berlin.de/top/PNX/pnstd-4.7.1.pdf>
5. Van der Aalst Wil. Making Workflow: On the Application of Petri Nets to Business Process Management. Application and Theory of Petri Nets 2002, Springer-Verlag, Berlin, 2002. P. 1—19.
6. Van der Aalst Wil. The Application of Petri Nets to Workflow Management // The Journal of Curcuits, Systems and Computers, 1998. P. 21—66.
7. Rusienkiewicz M., Sheth A. Specification and execution of Transactional Workflow. Modern Database Systems: the Object model, Interoperability and Beyond, Addison-Wesley, 1994. P. 37—54.
8. Питерсон Дж. Теория сетей Петри и моделирование систем. — М.: Мир, 1984.
9. Филатов М. А. CASE-технология анализа потоков работ // Науч. сессия МИФИ-2004: Сб. науч. тр. в 15 т. — М.: МИФИ, 2004. Т. 14. С. 50.
10. Ермаков А. В., Филатов М. А. CASE-система ER&LC System 2.0 с поддержкой концепции ЖЦ сущностей // Науч. сессия МИФИ-2002. — М.: МИФИ, 2002. Т. 13. С. 11—13.

