

ИССЛЕДОВАНИЕ МЕТОДИКИ ПРОЕКТИРОВАНИЯ РЕЛЯЦИОННЫХ БАЗ ДАнных НА ОСНОВЕ СЕТЕВОЙ МОДЕЛИ

Исследована методика преобразования информации табличного вида в реляционные базы данных на основе использования сетевой модели. На основе анализа сетевой модели скорректирована операторная модель.

Задачами исследования, решение которых должны обеспечить модель и математический аппарат, являются выявление и исключение ситуаций, приводящих к закливанию процесса интерактивного взаимодействия, недостижимых ситуаций в процессе проектирования, а также анализ функциональной полноты проекта системы. При этом модель должна отражать асинхронные процессы, предоставлять возможность моделировать протекание параллельных процессов и абстрагироваться от разделения ролей между разработчиком БД, системой специальных содержательных вопросов и операциями ввода-вывода. В качестве аппарата формализации, полностью удовлетворяющего сформулированным требованиям, выбран аппарат сетей Петри [1]. Анализ, выполненный в работе [2], показал, что в данном случае для моделирования достаточно использовать аппарат классических сетей Петри.

Автор настоящей статьи разработал операторную модель автоматизированного преобразования информации табличного вида (ИТВ) к реляционной базе данных (РБД) и этому посвящена отдельная глава монографии, которая опубликована совместно с А.В.Балдиным [2]. Операторная модель приведена на рис. 1.

Начальный уровень представления интерактивного взаимодействия разработчика и системы в виде сети Петри построен на основе операторной модели (рис. 2).

Моделям ИТВ, РБД, системы оценок (СО), средств преобразования (СП) поставлены в соответствие положения сети $\{P\}$. Операторам преобразований и принятия решений (ОП и ОПР соответственно) поставлены в соответствие переходы сети $\{t\}$. Имена положений и переходов сети Петри соответствуют номерам компонентов операторной модели. Индексы в операторах модели “р”, “н”, “к”, “с” соответствуют операторам преобразования ИТВ к реляционным таблицам, операторам нормализации, назначения ключевых полей, назначения связей между таблицами. Индекс “и” используется для оператора импортирования.

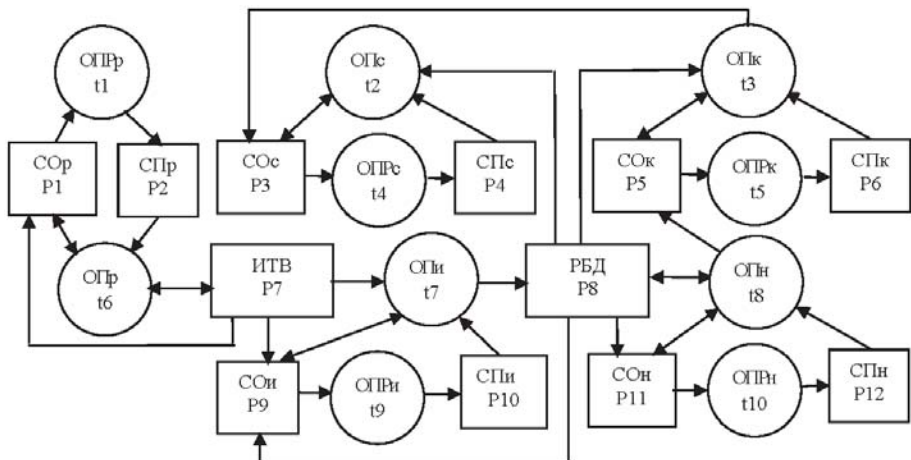


Рис. 1. Операторная модель преобразования ИТВ к РБД

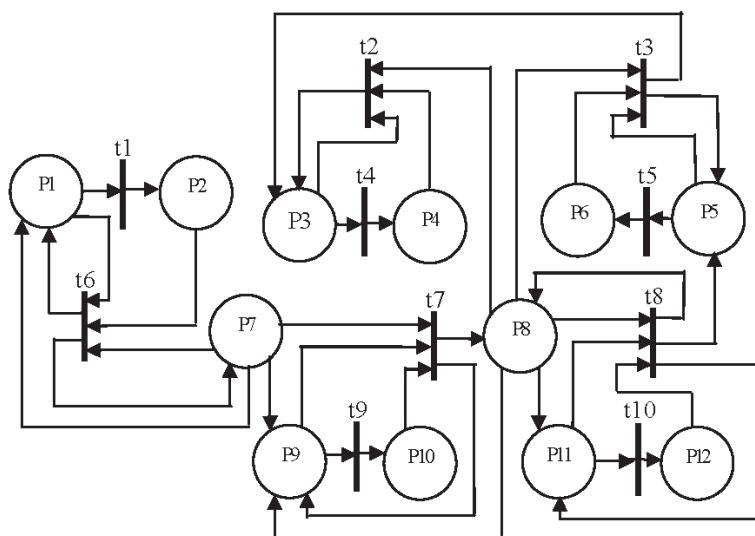


Рис. 2. Модель процесса преобразования ИТВ к РБД в виде сети Петри

Сеть Петри представляется тройкой $P = \langle P, T, P \rangle$, где P — множество положений; T — множество переходов; H — множество маркеров [3].

С помощью построенной модели можно выполнить анализ устойчивости сети, или анализ способности отражать реальные процессы интерактивного взаимодействия разработчика и автоматизированных средств проектирования в ходе проектирования РБД на основе существующей ИТВ.

Сеть Петри является устойчивой, если она имеет потоковое назначение $f_i > 0$ для каждого перехода, где потоковое назначение — это функция, которая приписывает каждому переходу положительное ра-

циональное число, называемое ее потоком. Потокое назначение для сети Петри должно удовлетворять следующим требованиям: каждая дуга переносит поток, равный потоку перехода, к которому присоединена эта дуга; для каждого положения сумма потоков входных дуг должна равняться сумме потоков выходных дуг [1].

Пусть для каждого перехода сети (см. рис. 2) назначен поток f_i . Для каждого положения P_i запишем уравнения потоков, которые не должны противоречить друг другу, если данная сеть устойчива. Уравнения сведены в табл. 1.

Таблица 1

Уравнения потоков

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	
P1	-f1					-f6 + f6					=0
P2	+f1					-f6					=0
P3		+f2-f2	+f3	-f4							=0
P4		-f2		+f4							=0
P5			+f3 - f3		-f5			+f8			=0
P6			- f3		+f5						=0
P7						-f6 + f6	-f7				=0
P8		-f2	-f3				+f7	-f8+f8			=0
P9							-f17 + f17		-f9		=0
P10							-f7		+f9		=0
P11								-f8 + f8		-f10	=0
P12								- f8		+f10	=0

Перепишем полученные уравнения в более компактной форме:

$$\begin{aligned}
 & -f_1 - f_6 + f_6 = 0; \\
 & +f_1 - f_6 = 0; \\
 & +f_2 - f_2 + f_3 - f_4 = 0; \\
 & -f_2 + f_4 = 0; \\
 & +f_3 - f_3 + f_8 - f_5 = 0; \\
 & -f_3 + f_5 = 0; \quad -f_7 - f_6 + f_6 = 0; \quad -f_3 + f_7 + f_8 - f_8 = 0; \\
 & -f_7 + f_7 - f_9 = -f_7 + f_9 = 0; \quad -f_8 + f_8 - f_{10} = 0; \quad -f_8 + f_{10} = 0.
 \end{aligned}$$

После преобразований уравнения примут вид

$$\begin{aligned}
 & f_1 = 0; \quad f_1 = f_6; \quad f_3 = f_4; \quad f_2 = f_4; \quad f_8 = f_5; \quad f_3 = f_5; \quad f_7 = 0; \\
 & f_7 = f_2 + f_3; \quad f_9 = 0; \quad f_7 = f_9; \quad f_{10} = 0; \quad f_8 = f_{10}.
 \end{aligned}$$

Выполнив соответствующие подстановки, получим:

$$f_1 = f_2 = f_3 = f_4 = f_5 = f_6 = f_7 = f_8 = f_9 = f_{10} = 0,$$

т.е. все потоковые значения нулевые. Это противоречит условию устойчивости системы.

Рассмотренная сеть не является устойчивой, а следовательно, процессы, ею описываемые, также являются неустойчивыми.

Поэтому модель нуждается в модификации, иначе методика, построенная на базе полученной модели, вероятно, будет содержать ошибки. Как видно из системы уравнений, нулевые потоковые значения, которые привели к нулевым потоковым значениям для всех положений, связаны с положениями P_1, P_7, P_9, P_{11} . Естественно предположить, что если исправить ситуацию для перечисленных положений, то она исправится и для сети в целом.

Вернемся к прообразам положений P_1, P_7, P_9, P_{11} и попытаемся выяснить сущность ошибки.

Положение P_1 отражает в модели (см. рис. 1) систему оценки (СОр) соответствия ИТВ требованиям, предъявляемым к реляционным таблицам. Из рис. 1 видно, что СОр по выходам связана с двумя операторами, а по входам — с одним (связь с ИТВ не является связью с оператором). Из связи такого рода можно сделать ложный вывод о том, что СОр получается непосредственно из ИТВ без всяких преобразований. На самом деле это не так. Необходимо выполнить ряд оценок, что сопряжено с действиями и, значит, с операторами. Поэтому в операторной модели преобразования введем соответствующий оператор ОСОр, а в сети Петри — переход t_{11} . Этот оператор и другие компоненты обновленной операторной модели отражены на рис. 3, который построен после выполнения всех необходимых манипуляций с сетью Петри.

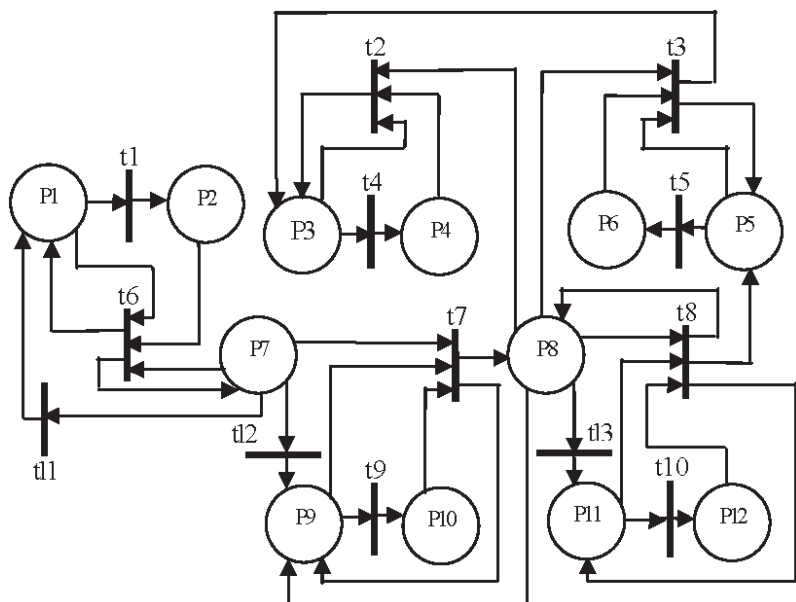


Рис. 3. Преобразованная сеть Петри

Прообразом положения P_7 является ИТВ, прообразом положения P_9 — система оценки импорта (СОи). Эти компоненты модели связаны между собой, поэтому вначале попытаемся нормализовать ситуацию с одной из компонент, а другую рассмотрим позже. Компонента СОи по входу связана с одним оператором ОПи, а по выходу — с двумя операторами ОПРи и ОПи. Имеется связь между ИТВ и СОи по входу, но она предполагает, что оценки необходимых действий по импорту выполняются без участия специальных средств. На самом деле это не так. В процессе импорта задействованы специализированные средства СУБД, использование которых не отражено. В связи с этим в операторную модель (см. рис. 1) необходимо добавить дополнительный оператор ОСОи между ИТВ и СОи, а в модель на основе сети Петри добавим соответствующий переход t_{12} .

Прообразом положения P_{11} является система оценки нормализации (СОн) таблиц РБД. Компонента СОн по входу связана с одним оператором ОПн, а по выходу — с двумя операторами ОПРн и ОПн. Это и нарушает потоковое равновесие, а СОн связана по входу непосредственно с РБД. Если бы они были связаны посредством оператора, то фрагмент модели, связанной с СОн, не привносил бы отрицательную составляющую в устойчивость сети. Действительно, в операторной модели ошибочно отражено состояние процесса оценки соответствия таблиц РБД нормальным формам. Это, безусловно, действие и даже система действий, которая в операторной модели должна быть отражена хотя бы одним оператором. В связи с этим в операторную модель (см. рис. 1) добавим оператор ОСОн, а в модель, полученную на основе сети Петри, — соответствующий переход t_{13} .

В результате выполнения модификаций получим сеть Петри, приведенную на рис. 3.

Несмотря на выявление и исключение теперь уже очевидных ошибок, полной уверенности в том, что полученная сеть стала устойчивой, нет. Дело в том, что в сеть добавлены три новых перехода. Тем самым изменено потоковое назначение сети не только для рассмотренных положений, но и для положений, связанных с ними через переходы.

Автором выполнена серия итераций исключения противоречий в моделях. Конечный результат отражен на рис. 4.

В результате выполненных преобразований исходной сети Петри получена устойчивая сеть. Функционирование такой сети не приведет к коллизиям перемещения информационных потоков. Поскольку полученная сеть соответствует операторной модели (рис. 5), то модель методики преобразования ИТВ к РБД является устойчивой, что исключает ошибки в методике, связанные с нарушением баланса информационных потоков, на ранних этапах ее разработки и реализации.

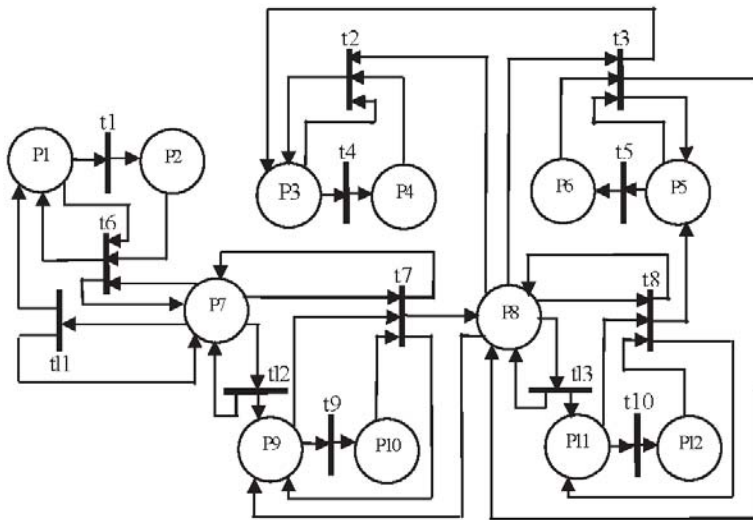


Рис. 4. Конечная сетевая модель

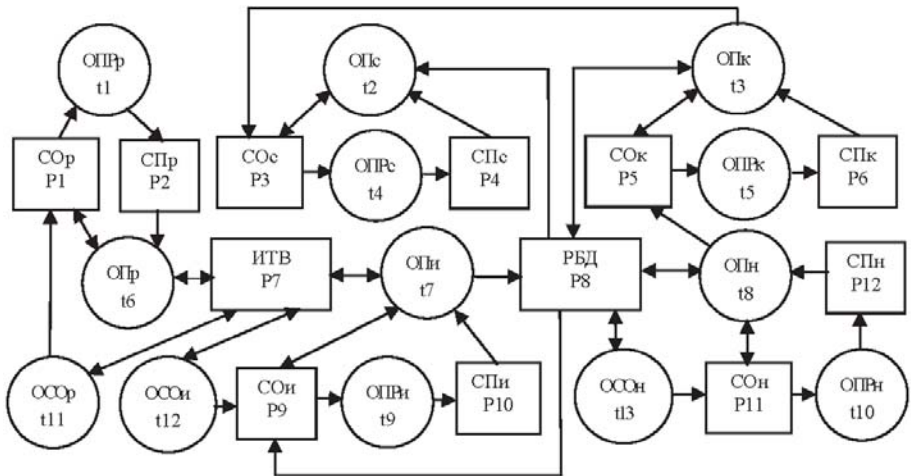


Рис. 5. Конечная операторная модель

Может создаться ложное впечатление о том, что операторная модель подгоняется под требования используемого математического аппарата. На самом деле наоборот, математический аппарат сетей Петри позволяет выявить коллизии в операторной модели, которая построена эмпирическим путем и не гарантирована от ошибок.

В результате выполненных преобразований, с одной стороны, получена операторная модель методики преобразования ИТВ к РБД, свободная от принципиальных ошибок, а с другой стороны, получено формальное описание этой модели, которое можно использовать для ее дальнейшего исследования и улучшения. В связи с тем, что в формальном описании модели выявлены и исключены принци-

альные ошибки, оно является адекватным реальному моделируемому процессу.

Операторная модель позволяет сделать выводы о составе операторов (методов, алгоритмов и средств), которые необходимо разработать для реализации полнофункциональной системы проектирования; о составе системы оценок или функций (методов, алгоритмов и средств), которые необходимо разработать для реализации полнофункциональной системы проектирования; о связях между методами и средствами в разрабатываемой системе.

Кроме того, операторная модель иллюстрирует порядок и правила использования методов и средств, задействованных в процессе решения задач преобразования ИТВ в РБД в разрабатываемой системе.

Что касается динамики функционирования сети Петри и соответствующей ей операторной модели, то она в данном случае отражена лишь отчасти и сводится к ориентации информационных потоков.

Их можно визуально отслеживать и делать необходимые выводы. Однако для более глубокого анализа динамических свойств модели и соответствующего объекта исследования оправданно использование специальных средств – маркерных сетей Петри, которые рассмотрены далее.

Динамику функционирования системы можно моделировать перемещением маркеров в сети в соответствии с правилами перехода

$$M'(P) = M(P) - P(t_i) + H(t_i),$$

где $M(P) = (M(P_1), M(P_2), \dots, M(P_N))$ – разметка сети [1].

Инициация всего процесса преобразования начинается с ИТВ и осуществляется разработчиком. Кроме того, при достижении положения P_8 , соответствующего РБД в операторной модели, разработчик может инициировать процесс и в положении P_7 . На сетевой модели это отображается маркером, который изначально располагается в P_7 . По мере необходимости в P_7 может добавляться маркер. При достижении маркером положения P_8 в это положение может быть при необходимости добавлен маркер. Это отражает реальные prerogatives разработчика. Собственно процесс моделирования функционирования системы посредством маркеров состоит в следующем. Маркеры перемещаются по сети из положения в положение. При этом маркеры могут быть перемещены лишь в том случае, если сработает входной переход, а он срабатывает, когда во всех положениях, которые связаны с ним по выходу, имеются маркеры. В качестве примера рассмотрим рис. 6.

Переход t_1 не сработает, пока в положении P_1 и P_2 не окажутся маркеры. Когда маркеры попадают в положения P_1 и P_2 , возможно срабатывание перехода t_1 . При этом маркеры из положений P_1 и P_2

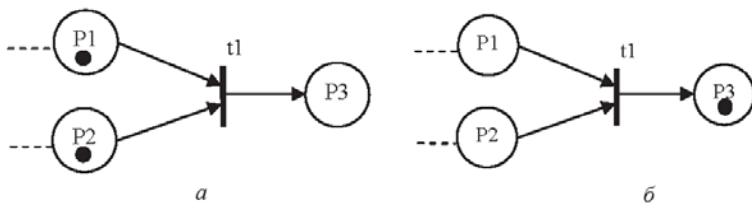


Рис. 6. Иллюстрация срабатывания перехода

удаляются и один маркер размещается в P_3 , что и проиллюстрировано на рис. 6, *а* и *б*.

Если данное правило поставить в соответствие реальным процессам, которые моделируются, то в нашем случае оператор выполняется лишь тогда, когда все необходимые данные имеются в наличии.

Перемещение маркеров по сети позволит сделать заключение о том, является ли сеть живой, т.е. существует ли принципиальная возможность срабатывания всех переходов. Кроме того, перемещение маркеров позволит сделать вывод о том, является ли сеть достижимой, т.е. достижимы ли все положения сети.

Для визуального отображения перемещения маркеров по сети с демонстрацией всех возможных состояний сети потребовалось около 50 рисунков, подобных рис. 4, только с различной разметкой положений. В связи со значительным объемом информации, представленной в виде рисунков, и плохой обзорности этой информации воспользуемся представлением разметки сети в виде кортежей. В этом случае порядковый номер в кортеже соответствует числу маркеров в соответствующем положении. Начальной маркировке сети соответствует следующая разметка сети:

$$M_1 = \{000000100000\}.$$

Срабатывание перехода t_{11} приведет к очередной разметке сети:

$$M_2 = \{100000100000\}.$$

Для удобства записи разметок сети и упрощения последующего их анализа последовательность разметок представлена в виде таблицы, фрагмент которой приведен в табл. 2. Возможные маркировки сети представлены в виде таблицы из 54 строк.

После просмотра последней строки таблицы сделан вывод о том, что все переходы сработали, причем многие из них неоднократно. Это свидетельствует о том, что полученная сеть живая, а значит, и процесс соответствующей сети живой, т.е. имеется принципиальная возможность выполнения всех операторов модели методики и достижимы все состояния моделируемой человеко-машинной системы.

Таким образом, анализируемая сеть Петри является достижимой, а значит, и все состояния соответствующего объекта исследования

Возможные маркировки сети

N	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	T
1	0	0	0	0	0	0	1	0	0	0	0	0	t0
2	1	0	0	0	0	0	1	0	0	0	0	0	t11
3	0	1	0	0	0	0	1	0	0	0	0	0	t1
4	1	1	0	0	0	1	0	0	0	0	0	0	t11
5	1	0	0	0	0	0	0	0	0	0	0	0	t6
6	0	1	0	0	0	0	0	0	0	0	0	0	t1
7	0	1	0	0	0	0	1	0	0	0	0	0	t0

достижимы исходя из логики и динамики функционирования модели процесса преобразования ИТВ в РБД.

В принципе процесс формирования очередных маркировок может осуществляться до бесконечности, так как все положения сети имеют непосредственную или опосредованную обратную связь.

В реальной ситуации процесс завершается при достижении необходимого результата. Принципиальная возможность срабатывания всех переходов свидетельствует о еще одном свойстве сети – отсутствии тупиковых положений. Это, в свою очередь, указывает на то, что в объекте моделирования отсутствуют тупиковые состояния.

Анализируя таблицу, выявили, что сеть является безопасной, т.е. выполняется условие, при котором в любой момент времени в каждом положении располагается не более одного маркера.

Из этого следует, что моделируемая система функционирует в стационарном режиме, т.е. при функционировании системы ни для одной из ее проектных процедур невозможна ситуация, когда формируется очередь на обработку входных данных, что исключает “зависание” процесса автоматизированного проектирования.

Резюмируя сказанное, можно сделать вывод о том, что разработанная сетевая модель устойчивая, живая, достижимая, безопасная, в ней отсутствуют тупиковые положения. Эта сеть соответствует реальному процессу преобразования ИТВ в РБД, и операторная модель преобразования обладает теми же свойствами.

Таким образом, операторная модель свободна от принципиальных ошибок и может быть использована в качестве основы для разработки и реализации методики проектирования РБД с использованием существующей ИТВ, а также в качестве основы для разработки соответствующих методов решения проектных задач, алгоритмов, используемых в рамках методов, и программных средств, которые в комплексе и составят методику.

СПИСОК ЛИТЕРАТУРЫ

1. Горбатов В. А. Основы дискретной математики. – М.: Высш. шк., 1986. – 311 с.
2. Балдин А. В., Брешенков А. В. Анализ проблемы проектирования реляционных баз данных на основе использования информации табличного вида и разработка модели методики проектирования. – М.: Изд-во МГТУ им. Н.Э. Баумана, 2007. – 150 с.
3. Кузнецов О. П., Адельсон-Вельский Г. М. Дискретная математика для инженера. – 2-е изд. перераб. и доп. – М.: Энергоатомиздат, 1988. – 480 с.

Статья поступила в редакцию 23.01.2007

Александр Владимирович Брешенков родился в 1955 г., окончил МВТУ им. Н.Э. Баумана в 1982 г. Канд. техн. наук, доцент кафедры “Компьютерные системы, комплексы и сети” МГТУ им. Н.Э. Баумана. Автор 70 научных работ в области САПР ЭВМ и баз данных.



A.V. Breshenkov (b. 1955) graduated from the Bauman Moscow Higher Technical School in 1982. D. Sc. (Eng.), assoc. professor of "Computer Systems, Complexes and Networks" department of the Bauman Moscow State Technical University. Author of 70 publications in the field of systems of automated design and data bases.

УДК 681.3.01

Б. И. Рабинович

РЕДАКТОР ШАБЛОНОВ СОЕДИНЕНИЙ КАК СРЕДСТВО ИНТЕГРАЦИИ БАЗЫ ЗНАНИЙ СИСТЕМЫ “АНАЛИТИК” С ВНЕШНИМИ ИСТОЧНИКАМИ ДАННЫХ

Рассмотрены методы, позволяющие расширить возможности логико-аналитической системы “Аналитик” в плане хранения и обработки информации. В качестве хранилища структур знаний предложено использовать систему управления базами данных Oracle, которая допускает работу с большими объемами данных. Рассмотрены варианты подключения к системе “Аналитик” в качестве дополнительных источников информации внешних баз данных, что позволит пользователю получать более полную информацию об интересующем объекте.

Одной из актуальных проблем взаимодействия информационных потоков является обработка больших потоков неформализованной информации — текстов естественного языка. Для этого в рамках плановых тем Института проблем информатики РАН были разработаны