

## Представление объектов и знаний в информационной среде «Система анализа форм отчетности»

С. Н. Богомолов, Ахмад Ибрахим, П. А. Иващенко  
*Харьковский национальный университет им. В.Н. Каразина, Украина*

The algorithm of performance of the arbitrary form of the reporting by the way systems of trees is offered. Is considered frames performance of the making forms of the reporting. The algorithm of creation of the production rules in to the «System of the analysis of the forms of the reporting» is given.

### 1. Введение

Создание и эксплуатация форм отчетности, адекватно и полно отражающих социально-экономическую действительность, – это проблема, которая всегда является актуальной для предприятий, учреждений любой страны. Формы отчетности (ФО) постоянно совершенствуются, приближаясь к международным стандартам [1, 2]. Проблема заключается в обеспечении возможности выявления направлений подобного совершенствования. Одно из направлений ее решения состоит в построении и использовании соответствующего инструментария. По нашему мнению, он должен включать в себя:

аналитическую составляющую, назначение которой – обеспечить анализ адекватности и полноты отражения информации, представленной в ФО;

информационно-математическую составляющую, которая позволяет систему используемых в ФО показателей представить в виде информационных (фреймы) и математических объектов (лес деревьев), хранимых на физическом уровне списком;

графическую составляющую (редактор графов), обеспечивающую визуализацию информационных представлений ФО.

В целом такой инструментарий может характеризоваться как система поддержки принятия решений (СППР), ориентированная на анализ и проектирование форм отчетности. В статье рассматривается подход к алгоритмизации представления ФО в виде списков, а также фреймовый инструментарий, обеспечивающий формирование правил вывода.

### 2. Анализ достижений и публикаций по теме исследования данной проблемы.

ФО содержат социально-экономическую информацию, организованную специальным образом. В них количественно отражена фотография субъекта хозяйственной деятельности за определенный период. От качества и полноты информационного представления зависит правильность восприятия и качество

выводов и принимаемых решений вышестоящими органами управления [1, с.184; 2, с. 27].

Аналитическая составляющая СППР – это наименее формализованный блок, использующий информационные представления из информационно-математического блока. Рассмотрим подробнее существующие подходы к графовым представлениям информации.

Под графовой визуализацией понимают такое представление сложной структуры, которое изображается, например, в виде набора точек (вершин графа) и системы связей между ними (дуг-ориентированных ребер). Автоматизация систем визуализации происходит в двух направлениях [5, с.414]. Первое направление связано со специализацией, ориентированной на графовые модели с определенной семантикой и топологией. Второе представлено универсальными системами обработки графовых моделей. Примерами могут служить программы Graph, VIZ и др.<sup>1</sup>. Для наших целей важным при изображении иерархической графовой модели представляется возможность наглядного представления свойств, которые могут быть связаны с двойственной ролью фрагмента (свойства типа «часть-целое»). Иначе говоря, в одной ситуации фрагмент может быть графом, в другой – вершиной.

Представляется актуальной разработка адекватных методов информационного моделирования предметной области системы анализа форм отчетности с помощью теории фреймов.

Несмотря на наличие большого количества приложений концепции фреймов к решению задач искусственного интеллекта и, в частности, к экспертным системам, формального определения фрейма не существует [6, с. 113]. В основу каждой из систем искусственного интеллекта закладывают собственную информационную структуру для реализации базовых концепций. Она базируется на представлении знаний с помощью фреймовой модели. Однако подходы, сходные с фреймовой моделью, развиваются. Если объединить концепцию фреймов с указанными подходами, то можно разработать законченную теорию построения информационных моделей, соответствующих предметной области системы анализа форм отчетности (САФО).

Трудностью представления знаний и языка фреймов является отсутствие формальной семантики. Это затрудняет сравнение свойств представления знаний различных языков фреймов, а также полное логическое объяснение языка фреймов.

Целью настоящей статьи является исследование возможности представления *любой* формы отчетности (документа) фреймовыми конструкциями для получения информации о качестве, как самой формы отчетности, так и о точности (безошибочности) ее заполнения.

Для модели представления знаний в САФО предлагается использовать продукционно-фреймовый формализм [6, с. 235-269; 7, с. 324-330; 8, с. 68-75]. Выбор и разработка такой гибридной модели представления знаний были осуществлены в результате исследования объекта (формы отчетности), анализа

---

<http://www.math.nsc.ru/LBRT/k3/Graph/Instruc.html>

состава и характера задач оценки эффективности (информативности) форм отчетности (ФО) [3, 4].

### 3. Постановка задачи

Задача состоит, во-первых, в построении алгоритма, реализующего преобразование представления ФО как части электронной таблицы с адресом

<адрес\_начального\_столбца\_адрес\_начальной\_строки :  
адрес\_последнего\_столбца\_адрес\_последней\_строки>

в виде набора списков. Каждый элемент списка – это ветвь дерева, ассоциированная с показателем. Показателем в данной ситуации выступает информационная единица, состоящая из набора реквизитов-признаков и одного реквизита-основания. Во-вторых, в разработке фреймовых конструкций ФО, ориентированных на получение информации о качестве как самой ФО, так и о точности (безошибочности) ее заполнения.

### 3. Основной материал

Форма отчетности представляет собой документ, содержащий заголовок, тело (таблицу) ФО и заключительную часть, которая представлена атрибутами руководителя и лица, отвечающего за подготовку ФО (главный бухгалтер, начальник планово-финансового отдела и т.п.). ФО может быть представлена на бумажном носителе и в электронном виде (в электронной таблице).

Как объект электронной таблицы ФО представляет собой совокупность ячеек. Некоторые из них объединены.

Поскольку предполагается анализировать любую ФО (из состава действующих), то считаем, что структура подлежащего (включая число строк) и структура сказуемого (включая число столбцов) заранее неизвестны. Наличие блока объединенных ячеек учитывается адресом его левой верхней ячейки.

Рассмотрим форматы и структуры данных, используемых при алгоритмизации и программировании задачи построения леса деревьев формы отчетности.

По содержанию ФО ее анализ может быть выполнен как по отдельным ячейкам, так и по совокупностям связанных ячеек. При этом могут учитываться не только строки и столбцы, но и их части.

Наряду со стандартным набором форматов ячеек, предусмотренных в Excel, данный подход требует использования нескольких форматов пользователя, которые учитывали бы такие особенности как незаполняемость, пустота ячеек и др.

Получающиеся в результате применения алгоритма связывания ячеек списки естественным образом могут быть объединены в группы по смысловым критериям. Совокупность списков однозначно характеризует лес деревьев.

Сформулируем детальнее требования к алгоритмам и программному обеспечению (ПО), используемым для анализа ФО.

ПО должно обеспечивать следующие возможности:

- добавлять связь между двумя ячейками (в нашей терминологии: вершинами);
- редактировать атрибуты (пользовательские форматы) ячеек;
- просматривать список путей, начинающихся в данной вершине.

Добавление вершин: после двух последовательных щелчков по двум ячейкам должна добавляться связь между вершинами, соответствующим этим ячейкам.

Результатом таких действий должно быть дерево, показывающее отношения между вершинами.

Редактирование атрибутов ячеек: для каждой ячейки необходимо, чтобы ПО позволяло изменять атрибуты вершины (например, вершина может соответствовать пустой ячейке, ячейке, содержащей число или текст).

Просмотр списка путей: необходимо, чтобы программа могла перечислять пути, начинающиеся в требуемой ячейке (т.е. собственно, строить дерево путей).

Авторами разработана надстройка над Excel, реализующая указанные требования.

Решение второй подзадачи состоит в следующем.

Аналогично [8] предлагается трехуровневая структура модели представления знаний (МПЗ). Первым (нижним) уровнем МПЗ выступает модель мира САФО, в котором отражены декларативные, экстенциональные и фактологические знания относительно проблемной области (альбом форм отчетности), пользователях системы (работниках бухгалтерий, планово-финансовых отделов и других служб предприятий, организаций, фирм). Возможность использования фреймового формализма в представлении знаний в модели мира САФО выражена в том, что вербальные знания групп профессионалов (экспертов) о структуре и строении предметной области допускают иерархическое, взаимосвязанное, вложенное описание (пример будет рассмотрен ниже).

Второй (средний) уровень МПЗ представлен базой правил (БП), в которой отражены знания процедурального и интенционального характера (методики заполнения ФО, характерные ошибки и др. информация). База правил должна быть структурирована в однозначном соответствии с составом предметной области (разделов ФО) и допускает представление в виде совокупности пакетов правил, являющихся независимыми продукционными системами.

Наконец, третий (верхний) уровень МПЗ содержит метазнания. Они отражают собственно характер и структуру знаний САФО. Метаправила позволяют осуществлять координацию работы с пакетами правил в БП. Метазнания позволяют облегчать процессы создания, отладки и поддержания в актуальном состоянии базы знаний (БЗ) для предметной области САФО.

Предлагается определить МПЗ четверкой

$$B = \langle F, R, M, I \rangle,$$

где  $F$  – модель мира САФО, представленная множеством фреймов ( $F = \{f_i, i = 1, \dots, Z\}$ );

$R$  – база правил;

$M$  – совокупность метазнаний;

$I$  – механизм вывода.

Каждый фрейм  $f_i$  определяется как пара:

$$f_i = \langle N_i^f, S_i^f \rangle,$$

где  $N_i^f$  – имя фрейма  $f_i$ ;

$S_i^f$  – слоты фрейма  $f_i$ .

Слоты фрейма  $f_i$  представлены множеством

$$S_i^f = \{(S_j^{f_i}), j = 1, \dots, J\}.$$

Каждый слот – это тройка:

$$S_j^{f_i} = \langle L_j^{f_i}, V_j^{f_i}, U_j^{f_i} \rangle,$$

где  $L_j^{f_i}$ , – имя слота  $S_j^{f_i}$ ;

$V_j^{f_i}$ , – значение слота  $S_j^{f_i}$ ;

$U_j^{f_i}$  – величина достоверности значения слота  $S_j^{f_i}$ .

База правил представлена множеством пакетов правил

$$R = (r_k), k = 1, \dots, K.$$

Каждое правило  $r_k$  – это пятерка

$$r_k = \langle N_k^r, A_k^r, K_k^r, W_k^r, E_k^r \rangle,$$

где  $N_k^r$  – имя правила  $r_k$ ;

$A_k^r$  – антецедент правила  $r_k$ ;

$K_k^r$  – консеквент правила  $r_k$ ;

$W_k^r$  – вес правила  $r_k$ ;

$E_k^r$  – трактовка (интерпретация, объяснение) правила  $r_k$ .

Антецедент правила  $r_k$  – это упорядоченное множество

$$A_k^r = \{(A_m^{r_k}), m = 1, \dots, M\},$$

где  $A_m^{r_k}$  – условный элемент антецедента.

Консеквент правила  $r_k$  – это упорядоченное множество

$$K_k^r = \{(K_n^{r_k}), n = 1, \dots, N\},$$

где  $K_n^{r_k}$  – действия в правой части правила.

Метазнания – это пара

$$M = \langle D^m, P^m \rangle,$$

где  $D^m$  – декларативные метазнания;

$P^m$  – процедуральные метазнания.

Декларативные метазнания – это множество

$$D^m = \{(F_p^m), p = 1, \dots, P\},$$

где  $F_p^m$  – метафрейм.

Процедуральные метазнания – это множество

$$P^m = \{(P_s^m), s = 1, \dots, S\},$$

где  $P_s^m$  – метаправило.

Механизм вывода – это множество способов рассуждений

$$I = \{(I_r), r = 1, \dots, R\},$$

где  $I_r$  – способ рассуждений.

Трехуровневая структура САФО может быть представлена рис. 1 (ср. с рис.11.2 из [8, с. 327]).

На рис. 1 приведен фрагмент БЗ из САФО. Первый уровень, как отмечено выше, содержит знания о структуре САФО, индивидуальные модели пользователей, совокупность решаемых задач и другую информацию. Объектами модели мира САФО являются формы отчетности, документы табличного типа, которые представлены в виде вложенных фреймов. В их слотах содержится информация фактографического и процедурного характера, описывающая объекты. Обычно несколько начальных слотов (а иногда и все) имеют жестко закрепленную семантику. С помощью некоторых слотов можно придать конкретную смысловую характеристику (интерпретацию) составляющих предметной области. Для ФО – это, в основном, древовидные, иерархические, родовидовые отношения между объектами.

Рассмотрим конкретный пример структуры базы знаний САФО относительно действующей ФО «Бухгалтерский баланс (Форма отчетности № 1)». Поскольку структура анализируемой ФО является типовой, достаточно рассмотреть лишь некоторые ее разделы.

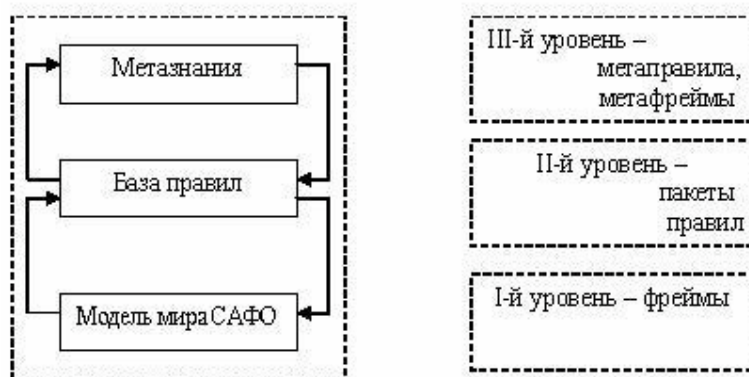


Рис. 1. Структура базы знаний САФО

Фрагмент реализации БЗ САФО представлен на рис. 2. На нем представлен раздел «Актив» «Бухгалтерского баланса» и часть раздела «Актив» – «Необоротные активы». С целью конкретизации базы правил в «Необоротных активах» представлены «Основные средства», состоящие из «Остаточной стоимости», «Начальной стоимости» и «Износа». Все записи относятся к началу отчетного периода.

Одной из задач САФО является интеллектуальный контроль правильности (корректности) заполнения ФО. Его суть, например, может состоять в системной проверке правильности заполнения ФО и расчетов. В нашем случае требуется проверить выполнение соотношения:

$$\text{Остаточная стоимость} = \text{Начальная стоимость} - \text{Износ}. \quad (1)$$

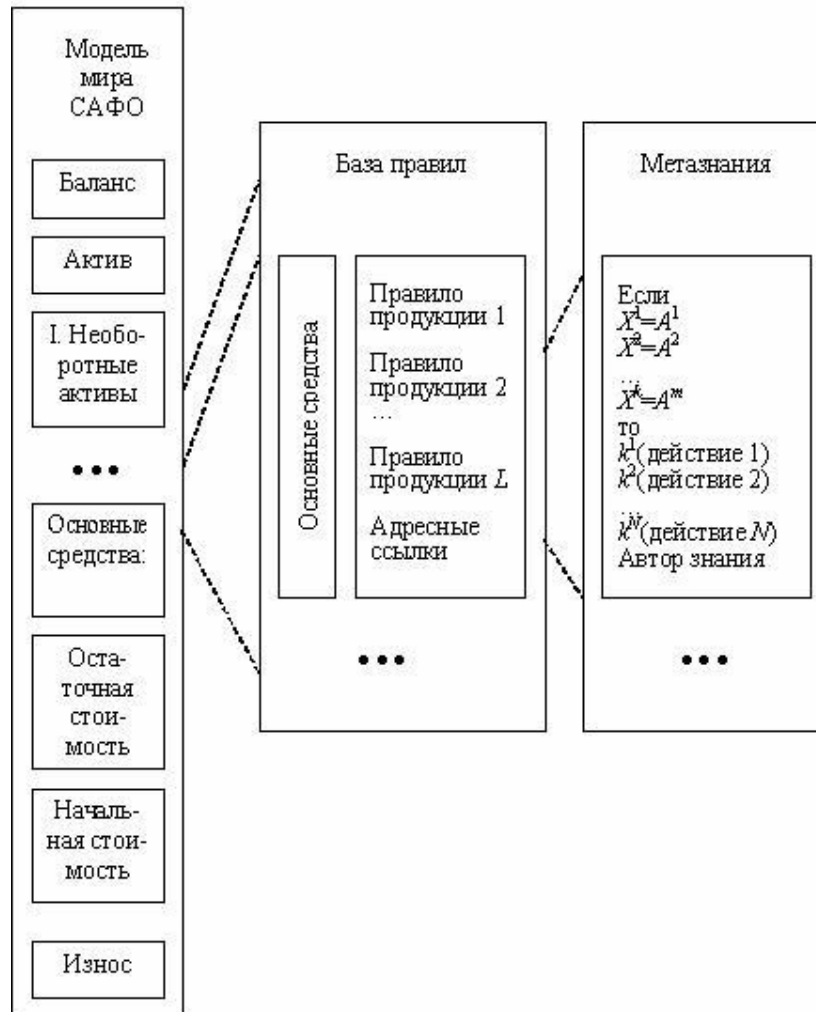


Рис. 2. Фрагмент реализации БЗ САФО

Интеллектуальность контроля правильности заполнения ФО состоит в предоставлении САФО возможных причин невыполнения соотношения (1) и выдаче рекомендаций по исправлению. Совокупность правил интеллектуального контроля может выглядеть следующим образом (в данном случае предполагаем единственность (1) в ФО):

Если соотношение (1) истинно, остановиться. Иначе:

1. Проверить правильность формулы (1). Если она верна, перейти к п. 2. Иначе рекомендовать ее исправить и затем перейти к п.1.

2. Проверить правильность формирования записи «Начальная стоимость» (Дебет счета 10). Если запись верна, перейти к п.3. Иначе рекомендовать

пользователю выполнить необходимые исправления. После корректировки перейти к п.2.

3. Проверить правильность формирования записи «Износ» (Кредит счета 131). Если запись верна, перейти к п.4. Иначе рекомендовать пользователю выполнить необходимые исправления. После корректировки перейти к п.3.

4. Выдать заключение о выполненных изменениях и остановиться.

Заметим, что многократная проверка в пп. 1-3 необходима для защиты от повторных ошибок.

Идеология рассмотренного алгоритма может быть легко распространена на всю ФО «Бухгалтерский баланс (форма отчетности № 1)».

На основе алгоритма вводятся предикаты, состояниям и действиям контролирующего алгоритма, формулируются правила вывода и стандартизируются. В результате образуется механизм логического вывода, который может быть использован в САФО в качестве инструментария аудита ФО.

#### **5. Выводы по данному исследованию и перспективы дальнейших исследований в данном направлении.**

Проведенное исследование показало возможность автоматизации процедур аудита форм отчетности в двух направлениях: аудит правильности заполнения и корректности построения ФО. Дальнейшие шаги исследования могут быть выполнены в направлении создания процедур интеллектуального распознавания ФО как по содержанию, так и по оформлению.

#### ЛИТЕРАТУРА

1. Парфенцева Н.О. Міжнародні статистичні класифікації в Україні. Впровадження й використання. – К.: Основи, 2000. – 351 с.
2. Методологічні положення зі статистики / Держ. Ком. Статистики України. – К.: ЗАТ „Август”, 2002. – Вип. I / редкол.: О.Г. Осауленко та ін. – 552 с.
3. Лазаренкова Г.М. Методологія оцінки вартості економічної інформації // Вісник ХНУ імені В.Н. Каразіна. Серія економічна. – № 580. – 2003. – С. 179-181.
4. Ахмад Ибрахим. Способ оценки эффективности информации форм отчетности // Вісник ХНУ імені В.Н. Каразіна. Серія економічна. – № 634. – 2003. – С. 107-110.
5. Касьянов В.Н., Евстигнеев В.А. Графы в программировании: обработка, визуализация и применение. – СПб.: БХВ-Петербург, 2003. – 1104 с.
6. Жолткевич Г.Н. Автоматизация проектирования технологической оснастки: теория и практика. К.: – Техніка, 1998. – 263 с.
7. Романов В.П. Интеллектуальные информационные системы. – М.: Экзамен, 2003. – 496 с.
8. Абдикеев Н.М. Проектирование интеллектуальных систем в экономике. – М.: Экзамен, 2004. – 528 с.