

10. Wooldridge M., van der Hoek W. On obligations and normative ability: Towards a logical analysis of the social contract // Journal of Applied Logic. – 2005. – Vol. 3. – P. 396-420.
11. Horty J.F. Agency and Deontic logic. – Oxford: Oxford University Press, 2001.
12. Boella G., van der Torre L. Substantive and Procedural Norms in Normative Multiagent Systems // Journal of Applied Logic. – 2008. Vol. 6. – P. 152-171.
13. Castelfranchi C. Modeling social action for AI agents // Artificial Intelligence. – 1998. – Vol. 103, № 1-2. – P. 157-182.
14. Felicissimo C., de Lucena C.J. P., Briot J.-P. A Norm-Based Approach for the Modeling of Open Multiagent Systems // Int. Conf. on Agents and Artificial Intelligence (Porto, Portugal, ICAART'09). – P. 540-546.
15. Garcia-Camino A., Noriega P., Rodriguez-Aguilar J.A. Implementing norms in electronic institutions // Proceedings of the 4th international joint conference on Autonomous agents and multiagent systems (The Netherlands). – 2005. – P. 667-673.

Статью рекомендовал к опубликованию д.т.н., профессор В.Н. Марков.

Миков Александр Иванович – Федеральное государственное бюджетное образовательное учреждение высшего профессионального образования «Кубанский государственный университет»; e-mail: alexander_mikov@mail.ru; 340050, г. Краснодар, ул. Ставропольская, 149; тел.: 89183456364; кафедра вычислительных технологий; д.ф.-м.н.; профессор; зав. кафедрой.

Mikov Alexander Ivanovich – Federal State Educational Establishment of Higher Professional Education “Kuban State University”; e-mail: alexander_mikov@mail.ru; 149, Stavropolskaya street, Krasnodar, 340050, Russia; phone: +79183456364; the department of computer technologies; dr. of phis.-math. sc.; professor; head of department.

УДК 004.652.4(045)

Ю.В. Полищук, Т.А. Черных

ФОРМИРОВАНИЕ ИНФОРМАЦИОННЫХ СИСТЕМ БОЛЬШИХ И СЛОЖНЫХ ТЕХНИЧЕСКИХ ОБЪЕКТОВ НА ОСНОВЕ КВАЗИСТРУКТУРИРОВАННЫХ МОДЕЛЕЙ ИНФОРМАЦИОННОГО НАПОЛНЕНИЯ

Рассмотрены основные задачи формирования информационных систем больших и сложных технических объектов на основе квазиструктурированных моделей информационного наполнения и предложены пути их решения. Применение квазиструктурированных моделей информационного наполнения реализует единый интерфейс доступа к фактографическим данным сопутствующего эксплуатационного контента. Обработка последнего необходима при принятии управленческих решений для больших и сложных технических объектов. В контексте данной работы рассмотрена математическая модель информационного наполнения документа и алгоритм синтеза моделей данного вида. Алгоритм синтеза моделей базируется на использовании пяти параметров и построенной на их основе целевой функции, применение которой позволяет проводить анализ квазиструктурированных моделей с позиции эффективности описания информационного наполнения документа. Основным преимуществом использования информационных систем больших и сложных технических объектов на основе квазиструктурированных моделей информационного наполнения является возможность обработки фактографического контента без извлечения из документа-первоисточника, что способствует снижению количества ошибок при обработке данных.

Моделирование документов; обработка электронных документов; квазиструктурированная информация; автоматизированные информационные системы.

Y.V. Polishuk, T.A. Chernykh

FORMATION OF INFORMATION SYSTEMS LARGE AND COMPLEX TECHNICAL OBJECTS BASED ON MODELS SEMISTRUCTURED INFORMATION CONTENT

The technique of creating models semistructured content of electronic documents. Using semi-structured models implements the generic interface to the factual content of the document that allows you to automate the processing of the document content and use the information stored in electronic documents, management processes. Application models semistructured content provides a single interface to access factual data accompanying operational content. Processing of the last needed for management decisions for large and complex technical objects. In the context of this paper the mathematical model of document content and synthesis algorithm for this type of models. Synthesis algorithm is based on models using five parameters and constructed on the basis of their objective function, the use of which allows the analysis of semistructured models with efficiency position description content of the document. The main advantage of the use of information systems of large and complex technical objects on the basis of semistructured content model is the ability to handle factual content without removing the primary source of the document, thereby reducing the number of errors in data processing.

The modeling of documents; processing of electronic documents; the semistructured data; automated information systems.

Процесс эксплуатации больших и сложных технических объектов требует применения информационной системы, которая реализует хранение и эффективную обработку сопутствующей эксплуатационной информации, которая необходима для принятия управленческих решений.

В контексте данной работы под большими техническими объектами будем понимать объекты, имеющие развитую иерархическую структуру, под сложными техническими объектами будем понимать объекты, которые не могут быть разбиты на подсистемы.

Информация об объекте поступает в информационную систему двумя способами: в виде формализованных фактографических данных, получаемых от SCADA-систем, и в виде документов с квазиструктурированным информационным наполнением [1].

Среди документов единого вида можно выделить общую структуру, но оформление и порядок размещения информации в них будет различен. В этом случае мы имеем дело с квазиструктурированной информацией. Под квазиструктурированной информацией понимают информацию, в которой можно выделить некую структуру, однако структура эта заранее целиком или частично неизвестна либо может меняться с течением времени [2].

Перспективным направлением исследований является реализация единого интерфейса доступа к формализованным и к квазиструктурированным данным, расположенным непосредственно в электронном документе.

Для реализации данного интерфейса авторами в работе [3] была предложена математическая модель информационного наполнения документа:

$$S = \langle root, sObj, LObj, minOccurs, maxOccurs, sMet, Obj_smet \rangle, \quad (1)$$

где *root* – корневой объект, $root \in sObj$; *sObj* – конечное множество объектов, каждый из которых содержит фрагмент информационного наполнения документа (текст, рисунок и т.д.) или выполняет роль контейнера для одного или нескольких объектов.

Для объектов-контейнеров доступны следующие метасвойства: $smet_c$ – определяет объект в качестве контейнера; $mixed$ – разрешает использование объектов-потомков в произвольном порядке. $LObj$ – отображение, определенное на множестве $sObj$, такое что $sObj \xrightarrow{LObj} \{obj_1, \dots, obj_n\}$, где $obj_i \in sObj$ – дочерний объект; n – количество дочерних объектов; Obj_met – отображение, определенное на множестве $sObj$, такое что $sObj \xrightarrow{Obj_met} \{smet_c / smet_c, mixed / smet_1, \dots, smet_k\}$, где $smet_i \in sMet$ – метасвойство ограничения на содержимое объекта; $minOccurs$ – функция, определяющая минимально возможное количество раз использования объекта в модели; $maxOccurs$ – функция, определяющая максимально возможное количество раз использования объекта в модели.

Рассмотрим графическое представление электронного документа (рис. 1). Документ, представленный на рис. 1, состоит из пяти объектов. Объект A – выполняет роль контейнера для объектов B и C , объект B – выполняет роль контейнера для объектов D и E . Объекты A, B, D – обязательно должны быть использованы при разработке документа, объект C – является необязательным к использованию, объект E в рассматриваемом примере должен быть использован от трех до пяти раз. Объекту-контейнеру A соответствует метасвойство ограничения $smet_c$, а для объекта-контейнера B определено дополнительно метасвойство $mixed$. Объект C представлен числовым наполнением, т.е. ему соответствует метасвойство ограничения $smet_2$. Объекты D, E – имеют символьное информационное наполнение, которому соответствует метасвойство ограничения $smet_1$.

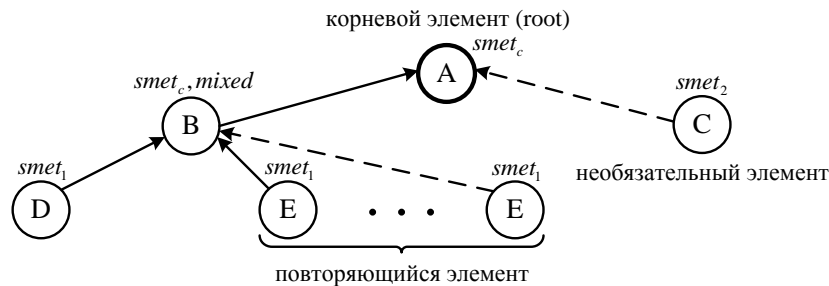


Рис. 1. Граф информационного наполнения документа

Документ может быть представлен с помощью модели следующим образом:
 $root = \{A\}$; $sObj = \{A, B, C, D, E\}$; $LObj(A) = \{B, C\}$, $LObj(B) = \{D, E \mid E, D\}$,
 $LObj(C) = \{\}$, $LObj(D) = \{\}$, $LObj(E) = \{\}$; $Obj_smet(A) = \{smet_c\}$,
 $Obj_smet(B) = \{smet_c, mixed\}$, $Obj_smet(C) = \{smet_2\}$,
 $Obj_smet(D) = \{smet_1\}$, $Obj_smet(E) = \{smet_1\}$; $minOccurs(A) = 1$,
 $maxOccurs(A) = 1$; $minOccurs(B) = 1$, $maxOccurs(B) = 1$;
 $minOccurs(C) = 0$, $maxOccurs(C) = 1$; $minOccurs(D) = 1$,
 $maxOccurs(D) = 1$; $minOccurs(E) = 3$, $maxOccurs(E) = 5$.

Применение единой квазиструктурированной модели информационного наполнения к коллекции однотипных документов реализует маркировку фактографических данных, расположенных непосредственно в документах.

Фрагмент электронного документа с информацией о результатах выполненных гидродинамических исследований на газовой скважине формата MS Word с размеченными фактографическими данными изображен на рис. 2.

Выполненный комплекс ГИС:								
Дата	Вид исследования	Режим исследования	Время замера	Интервал исследования, м	Руст, МПа Нач.	МПа Кон.	Рзатр, МПа Нач.	МПа Кон.
« RunCom « Data 04.09 Data »	« VidIsl ГК VidIsl »	« RegIsl стат ика длительная RegIsl »	TimeBZam 2 1:20 TimeEZam 2 2:50	« IntIsl 1250- 1560 IntIsl »	« DavUs 0 DavUstB	« DavUst 0 DavUstE	« DavZat 0 DavZatB	« DavZat 0 DavZatE RunCom
Тип прибора		Дата эталонировки	Начальник партии					
« Pribor РОН № 18 TypeP »		« DataEt 05.07.2006 DataEt »	DirPart СКОКОВ В.Ю. Pribor »					

Результаты ГИС:

« Result
Res По данным нейтрон-нейтронного каротажа газопыльчатые скопления за НКТ под пакером отмечаются в интервале 1303.8-1346 м, в интервале 1538-1556 м - газопыльчатый "пузырь". По сравнению с замером от 22.08.06 их положение не изменилось (см. прил.)»

Рис. 2. Фрагмент электронного документа MS Word с размеченными фактографическими данными

Таким образом, после разметки содержимого электронного документа и загрузки его в информационную систему доступ к фактографическому содержанию реализуется в автоматизированном режиме.

Для формализации процесса создания квазиструктурированных моделей информационного наполнения авторами был разработан алгоритм синтеза моделей данного вида, который подробно описан в работе [4]. Алгоритм синтеза моделей базируется на использовании пяти параметров и построенной на их основе целевой функции, применение которой позволяет проводить анализ квазиструктурированных моделей с позиции эффективности описания информационного наполнения документа.

В качестве первого параметра примем результат валидации документа. Рассматриваемый параметр характеризует соответствие модели стандарту на содержание рассматриваемых документов. При невыполнении условия валидации модель бракуется:

$$P_1 = \frac{Oobj(sObj_{DOC})}{Oobj(sObj_S)} = 1, \tag{2}$$

где $Oobj$ – функция, возвращающая количество обязательных объектов из множества, удовлетворяющих условию: $minOccurs(sObj_i) = 1$, где $sObj_i \in sObj$; $sObj_{DOC}$ и $sObj_S$ – соответственно множества объектов, использованных в документе, и объектов рассматриваемой модели. Данный параметр должен быть равен единице, так как использование обязательных параметров является условием применения модели документа.

Второй параметр характеризует степень детализации, т.е. размер контента использованных в документе объектов (структурных единиц) модели. Данный параметр должен быть минимизирован, так как при работе с объектами всегда проще укрупнять, чем детализировать информационное наполнение документа:

$$P_2 = \frac{\sum Len(sObj_i)}{Cobj(sObj_i)} \rightarrow \min, sObj_i \in sObj, \quad (3)$$

где $Cobj$ – функция, возвращающая количество объектов множества, удовлетворяющих условию $Obj_smet(sObj_i) \cap \{smet_c\} = \emptyset$.

Третий параметр характеризует плотность использования объектов модели, т.е. уровень проработки модели документа и характеризует равномерность распределения объектов (структурных единиц) модели по документу:

$$P_3 = \sqrt[k]{\prod_{j=1}^k Pc_j} \rightarrow \min, Pc_j = \frac{p_char_count(j)}{p_obj_count(j)}, \quad (4)$$

где Pc_j – плотность объектов модели на фрагменте документа (фрагмент равен странице документа или абзацу, т.е. документ состоит из k фрагментов); p_char_count – функция, возвращающая количество символов в определенном фрагменте; p_obj_count – функция, возвращающая количество объектов $sObj_i$ таких, что $Obj_smet(sObj_i) \cap \{smet_c\} = \emptyset$, целиком размещенных в заданном фрагменте. Рассматриваемый параметр должен быть минимизирован.

Четвертый параметр будет характеризовать насыщенность объектов (структурных единиц) модели в документе, т.е. качество описания информационного наполнения документа:

$$P_4 = \frac{char_count}{\sum Len(sObj_i)} \rightarrow \min, Obj_smet(sObj_i) \cap \{smet_c\} = \emptyset, \quad (5)$$

где $char_count$ – функция, возвращающая количество символов в документе; Len – функция, возвращающая количество символов из объекта. Данный параметр должен быть минимизирован, так как необходимо сократить неопиываемое информационное наполнение в документе.

Пятый параметр характеризует гибкость модели. Более гибкая модель позволяет эффективнее описать квазиструктурированное информационное наполнение документа, следовательно данный параметр должен быть минимизирован:

$$P_5 = \frac{Aobj(sObj_i) - (Uobj(sObj_i) + Robj(sObj_i))}{Aobj(sObj_i)} \rightarrow 0, \quad (6)$$

где $Aobj$ – функция, возвращающая количество объектов без учета корневого элемента $root$; $Uobj$ – функция, возвращающая количество объектов множества, удовлетворяющих условиям: $minOccurs(Obj_i) = 0$; $maxOccurs(Obj_i) = 1$; $Robj$ – функция, возвращающая количество объектов множества, удовлетворяющих условиям $minOccurs(Obj_i) = 1$; $maxOccurs(Obj_i) > 1$, где $sObj_i \in sObj$.

Для оценки качества рассматриваемых моделей введем скалярную критериальную целевую функцию (7). Ее значение для лучшей из рассматриваемых моделей будет минимальным:

$$R = \frac{\sum_{i=2}^5 P_i}{P_1} \rightarrow \min, \quad (7)$$

где P_i – значение i -го параметра анализируемой модели.

В случае, когда модель не соответствует информационному наполнению документа, т.е. не выполняется ее валидация, целевая функция (7) будет неопределена.

Для снижения трудоемкости процесса применения модели к документу было разработано программное средство «Программа разметки фактографического контента электронных документов с квазиструктурированным информационным наполнением на основе паттернов» [5], которое позволяет выполнять данный процесс в полуавтоматическом режиме. Основные принципы работы данной программы рассмотрены в работе [6].

Эффективность применения информационных систем на основе квазиструктурированных моделей информационного наполнения была подтверждена авторами в работе [7].

Таким образом, использование квазиструктурированных моделей информационного наполнения актуально при формировании информационных систем больших и сложных технических объектов, а основным преимуществом их использования является возможность обработки фактографического контента без извлечения из документа-первоисточника.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. *Полищук Ю.В.* Квазиструктурированный контент в управлении большими и сложными техническими системами. – Самара: СамГУПС, 2014. – 191 с.
2. *Палей Д.* Моделирование квазиструктурированных данных // Открытые системы. – 2002. – № 09. – С. 57-64.
3. *Полищук Ю.В., Черных Т.А.* Моделирование документов средствами Adobe FrameMaker // Вестник компьютерных и информационных технологий. – 2011. – № 5. – С. 35-40.
4. *Полищук Ю.В., Черных Т.А.* Синтез квазиструктурированных моделей информационного наполнения электронных документов // Вестник компьютерных и информационных технологий. – 2012. – № 6. – С. 20-27.
5. *Полищук Ю.В., Ларин А.В.* Программа разметки фактографического контента электронных документов с квазиструктурированным информационным наполнением на основе паттернов // Свидетельство о государственной регистрации программы для ЭВМ № 2013611328 от 9 января 2013 г.
6. *Полищук Ю.В., Ларин А.В.* Синтез квазиструктурированных моделей информационного наполнения электронных документов // Вестник компьютерных и информационных технологий. – 2013. – № 3. – С. 55-60.
7. *Полищук Ю.В., Черных Т.А.* Оценка эффективности автоматизированных систем управления на основе квазиструктурированного контента // Технологии разработки информационных систем (ТРИС-2013): материалы конференции. Т. 2. – Таганрог: Изд-во ТТИ ЮФУ, 2013. – С. 32-37.

Статью рекомендовал к опубликованию д.т.н., профессор И.Т. Ковриков.

Полищук Юрий Владимирович – Федеральное государственное образовательное учреждение высшего профессионального образования «Оренбургский государственный университет»; e-mail: Youra_Polishuk@bk.ru; 460018, г. Оренбург, просп. Победы, 13; тел.: 83532372534; кафедра компьютерной безопасности и математического обеспечения информационных систем; к.т.н.; доцент.

Черных Татьяна Александровна – e-mail: chatty84@mail.ru; тел.: 83532372537; кафедра информатики; к.т.н.; доцент.

Polishuk Yuri Vladimirovich – Federal State Educational Government-financed Institution of Higher Professional Education «Orenburg State University»; e-mail: Youra_Polishuk@bk.ru; 13, pr. Pobedy, Orenburg, 460018, Russia; phone: +73532372534; the department of mathematics and computer security, information systems; cand. of eng. sc.; associate professor.

Chernykh Tatiana Aleksandrovna – e-mail: chatty84@mail.ru; phone: +73532372537; the department of informatics; cand. of eng. sc.; associate professor.

УДК 004.4'22

А.Н. Беликов

ОБЗОР И АНАЛИЗ МЕТОДОВ И СРЕДСТВ ПРОЕКТИРОВАНИЯ КОНФИГУРИРУЕМЫХ ИНФОРМАЦИОННЫХ СИСТЕМ*

Рассматривается подход, основанный на использовании фреймворков, как платформы для разработки информационных систем. Данные платформы представляют собой динамические структуры. После конфигурирования динамические структуры становятся деятельностью со статической структурой – создаваемой информационной системой. Целью данной работы является анализ существующих CASE-средств на возможность проектировать конфигурирование информационной системы на базе фреймворка. Для этого, описываются основные свойства фреймворка. На их основе выделяются требования к CASE-средствам проектирования процесса конфигурирования информационных систем. Для того, чтобы провести анализ, существующие методы и CASE-средства проектирования условно разделяются на две группы: методы и средства структурного анализа и проектирования и методы и средства объектно-ориентированного проектирования. На основе проведенного анализа приводится концептуальная модель CASE-средства по проектированию конфигурирования информационных систем на базе фреймворка. Данная концептуальная модель CASE-средства учитывает: описание работ конечного пользователя и описание объекта (свойства объекта).

Проектирование; конфигурируемая информационная система; фреймворк; CASE-средство.

A.N. Belikov

REVIEW AND ANALYSIS OF METHODS AND TOOLS OF CONFIGURABLE INFORMATION SYSTEMS DESIGN

The paper presents an approach based on the use of frameworks as a platform for information systems development. These platforms are dynamic structures. After configuring the dynamic structures are activities with a static structure – i.e. they are the information system. The aim of this work is to analyze the existing CASE-tools on the criteria of opportunity to design the process of configuring the information system based on the framework. For this purpose, the main features of the framework are described. On this basis the requirements for CASE-design tools of configuring the information systems are marked up. In order to analyze the existing methods and CASE-design tools are conventionally divided into two groups: methods and tools of structural analysis and design techniques and tools for object-oriented design. Based on the analysis a conceptual model of CASE-tools for the configuration information systems design based on the framework is provided. This conceptual model of CASE-tool takes into account: the description of the end-user and a description of the object (object properties).

Design; configurable information system; framework; CASE- tool.

* Исследование выполнено при поддержке государственного задания 0110021005901621. Тема № 213.01-11/2014-17.